

UNCLASSIFIED



Integrated Sensing and Processing (ISP) Phase II: Demonstration and Evaluation for Distributed Sensor Networks and Missile Seeker Systems

Progress Report:

**4th Quarter Progress Report
1 December 2005 – 28 February 2006**

Acknowledgment of Support

This material is based upon work supported by the United States Air Force under Contract No. N00014-04-C-0437.

Contract No.: N00014-04-C-0437
Contract Line Item Number 0001
Deliverable Item: Publications (A001-004)

**Raytheon Company
P.O. Box 11337
Tucson, AZ 85734-1337**

DESTRUCTION NOTICE

For classified documents, follow the procedures in DOD 5220.22M, National Industrial Security Program Operating Manual (NISPOM), Chapter 5, Section 7, or DOD 5200.1-R, Information Security Program Regulation, Chapter IX. For unclassified, limited documents, destroy by any method that will prevent disclosure of contents or reconstruction of the document.

Distribution Statement A.
Approved for public release.
Distribution is unlimited.

© RAYTHEON MISSILE SYSTEMS (2005) UNPUBLISHED WORK

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227 7013 (Nov 1995)

UNCLASSIFIED

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 21 FEB 2006		2. REPORT TYPE		3. DATES COVERED	
4. TITLE AND SUBTITLE Fourth Quarterly Progress Report for Period of Performance			5a. CONTRACT NUMBER N00014-04-C-0437		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Harry Schmitt			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Raytheon Missile System,P.O. Box 11337,Tucson,AZ,85734			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The primary goal of this effort is to bring to maturity a select set of basic algorithms, hardware, and approaches developed under the Integrated Sensing and Processing (ISP) Phase I program, implement them on representative hardware, and demonstrate their performance in a realistic field environment. We have identified a few promising research thrusts investigated in ISP Phase I where field demonstrations are cost prohibitive but collected data sets are available. Here, we will conduct a thorough performance evaluation.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT 1	18. NUMBER OF PAGES 55	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

01 March 2006

Progress Report

**CDRL A001 No. 04
First Annual Report
and
Fourth Quarterly Progress Report for Period of Performance
1 December 2005 – 28 February 2006**

**Integrated Sensing Processor Phase 2
Program Manager: Dr. Harry A. Schmitt
Principal Investigator: Dr. Harry A. Schmitt**

Sponsored By:

**Defense Advanced Research Projects Agency/DSO
Dr. Carey Schwartz/DARPA DSO
Program Manager: Dr. Dan Purdy/ONR
Issued by ONR under Contract #N00014-04-C-0437**

Prepared By:

**Raytheon Systems Company
P.O. Box 11337
Tucson, AZ 85734**

EXECUTIVE SUMMARY

The primary goal of this effort is to bring to maturity a select set of basic algorithms, hardware, and approaches developed under the Integrated Sensing and Processing (ISP) Phase I program, implement them on representative hardware, and demonstrate their performance in a realistic field environment. We have identified a few promising research thrusts investigated in ISP Phase I where field demonstrations are cost prohibitive but collected data sets are available. Here, we will conduct a thorough performance evaluation.

TABLE OF CONTENTS

0. Technical Abstract	2
1. A. Program Summary	2
1. B. Program Status	3
1. C. Recent Accomplishments and Events	4
1. D. Near Term Events	4
2.0. Technical Progress and Accomplishments	5
2. A. Technical Progress	5
2.A.1. Raytheon Technical Progress	5
2.A.2. ASU Technical Progress	25
2.A.3. UM Technical Progress	26
2.A.4. FMAH Technical Progress	29
2.A.5. Georgia Tech Technical Progress	34
2.A.6. UniMelb Technical Progress	35
2. B. Publications	50
2. C. Conference Proceedings	50
2. D. Consultative and Advisor Functions	51
2. E. New Discoveries, Inventions or Patent Disclosures	51
2. F. Honors/Awards	51
2. G. Transitions	51
2. H. References	51
2. I. Acronyms	54

0. Technical Abstract

Advances in sensor technologies, computation devices, and algorithms have created enormous opportunities for significant performance improvements on the modern battlefield. Unfortunately, as information requirements grow, conventional network processing techniques require ever-increasing bandwidth between sensors and processors, as well as potentially exponentially complex methods for extracting information from the data. To raise the quality of data and classification results, minimize computation, power consumption, and cost, future systems will require that the sensing and computation be jointly engineered. ISP is a philosophy/methodology that eliminates the traditional separation between physical and algorithmic design. By leveraging our experience with numerous sensing modalities, processing techniques, and data reduction networks, we will develop ISP into an extensible and widely applicable paradigm. The improvements we intend to demonstrate here are applicable in a general sense; however, this program will focus on distributed sensor networks and missile seeker systems.

1.0. Management Overview and Summary

1. A. Program Summary

The Raytheon Company, Missile Systems (Raytheon) ISP Phase II program is a twenty-four month contract with a Period of Performance (PoP) covering 1 March 2005 to 28 February 2007. Raytheon has four universities and one small business as ISP Phase II subcontractors: Arizona State University (ASU); Fast Mathematical Algorithms and Hardware (FMAH); Georgia Institute of Technology (Georgia Tech); Melbourne University (UniMelb) and the University of Michigan (UM).

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

1. B. Program Status

The Raytheon ISP Phase II Program status can be summarized as remaining “on track.” All of the negotiations have been completed and all of the subcontractors are now under subcontract. However, we have incurred some schedule slips on both the distributed tracking and the Cooperative Analog Digital Signal Processing (CADSP) demonstrations. An updated schedule for the distributed tracking demonstration has been developed and is included in Section 2.0. While the revised schedule still supports a demonstration before 28 February, there is little room for further slippage. The current status of the CADSP imager hardware is discussed in Section 2.A.6.

The Program is still running below the spending plan; however, we expect to complete the contract on time and budget. As of 20 January 2006, 31% of contract funds had been expended with ~45% of the program complete. In part, the contract expenditure reflects an under-run due to delays in receiving invoices from our subcontractors. Raytheon has also under-run significantly. Initially the reduced Raytheon spending was to better align with the subcontractor schedules; however, we have also encountered difficulties with personnel availability. We believe that we have the Raytheon personnel availability issue resolved and should recover from the spending profile deviation.

One area of significant concern is the availability of a suitable radar test and integration engineer. We continue to work this issue, but have not been successful so far. Failure to resolve this problem soon is probably the highest risk for our program.

1. C. Personnel Associated/Supported

Raytheon

Dr. Harry A. Schmitt	Principal Investigator
Mr. Donald E. Waagen	Co-Principal Investigator
Dr. Sal Bellofiore	Distributed Sensing Lead
Mr. Thomas Stevens	Distributed Sensing Support
Dr. Robert Cramer	Mathematical Support
Mr. Craig Savage	Waveform Design and Control Lead
Dr. Nitesh Shah	High Dimensional Processing Data Lead

FMAH

Professor Paolo Barbano
Professor Ronald Coifman
Dr. Nicholas Coult

ASU

Professor Darryl Morrell
Professor Antonia Papandreou-Suppappola

Georgia Tech

Professor David Anderson
Professor Paul Hasler

UniMelb

Dr. Barbara LaScala

**ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)**

Professor William Moran
Dr. Darko Musicki
Dr. Sofia Suvorova

UM

Professor Al Hero
Dr. Neal Patwari

Significant Personnel Actions: There was one significant personnel change during the current PoP. Dr. Neal Patwari of the University of Michigan became actively involved in developing self localization approaches for the MICA-2/Z distributed tracking demonstration.

1. C. Recent Accomplishments and Events

In support of the distributed processing demonstrations and evaluations, Neal Patwari (UM) spent a week at Raytheon in November 2005 to discuss self localizations approaches for the nodes. Raytheon presented an overview of the mechanics of working with the Algorithms Verification Units (AVUs), installed the software necessary to program the AVUs and spent several days working with UM on implementation. Four of the AVUs were delivered to UM in December 2005. The remaining AVUs are available for distribution to other university personnel when needed.

An amended Technical Assistance Agreement (TAA) was approved by the U.S. State Department on 6 October 2005. The amended TAA expands the technical scope to cover the research areas added under the ISP Phase II program, adds two dual citizens at UniMelb, and also covers Raytheon, Australia. The amended TAA has been signed by UniMelb and is out for signature by the remaining parties.

Other Accomplishments and Events:

- Raytheon personnel (Waagen and Schmitt) visited Georgia Tech 16 November to discuss current hardware and algorithm status.
- Raytheon personnel (Waagen, Schmitt and Samuel) visited University of Maryland on 15 November to discuss waveform design research being conducted by Professor John Benedetto.
- Raytheon personnel (Waagen, Schmitt and Samuel) met with Professor Stuart Milner, Director of the Center for Networking of Infrastructure Sensors to discuss possible collaboration opportunities.
- Delivered report on mathematical foundation of FMAH waveform family design.

1. D. Near Term Events

- Dr. T.J. Klausutis (AFRL, Eglin) will visit Georgia Tech to discuss possible collaborative opportunities. In particular, there may be an opportunity to get the Georgia Tech CADSP imager included in AFRL test plans. AFRL, Eglin is particularly interested in learning more about the capabilities and maturity of the CADSP imager. As yet, we have been unable to confirm a date for this visit.
- Attend the DARPA Waveforms for Active Sensing (WAS) Program Review Meeting will be held March 14, 2006 in Portland, Oregon.

- Present “Cooperative Control of Multiple UAVs for Passive Geolocation,” at the Special Session on Cooperative Dynamic Systems, 2006 IEEE International Conference on Networking, Sensing and Control, Ft. Lauderdale, FL.

2.0. Technical Progress and Accomplishments

Again, most of the effort expended during the current PoP has been in the areas of data collection, sensor characterization, mathematical analyses, and algorithm development for the distributed tracking demonstration. Key technical focus areas include:

- The development, implementation and evaluation of accurate and scalable sensor self-localization approaches.
- Data collection for MICA-2/Z sensors characterization. The focus was on characterizing the acoustic sensors with the goal of developing a 1-bit on-mote detection algorithm.
- Preliminary characterization of the vibration sensor was continued.
- The development of distributed tracking algorithms at UniMelb. The refinement of the final tracking demonstration.

We have also expended significant effort in the following technical evaluation areas.

- Evaluation of High Dimensional Data Processing of collected field data.
- The investigation of mathematically rigorous approaches for the critical problem of handling out-of-sample extensions for High Dimensional Data Processing.
- Stochastic approaches for UAV control and passive geolocation.
- Algorithm definition and hardware development/test for the implementation and demonstration of the Georgia Tech CADSP imager.

Significant effort has gone into the development and refinement of a detailed test plan for the distributed tracking demonstration. While we still expect this to evolve some over time, the schedule shown in Table 1 is a pretty firm baseline. These technical focus areas are discussed in significantly more detail in Subsection 2.A, where the technical approaches for Raytheon and for each subcontractor are described. Preliminary experimental results are also summarized.

2. A. Technical Progress

2.A.1. Raytheon Technical Progress

2.A.1.a. Distributed Sensor Demonstration

Wireless low-power sensor networks have gained much deserved attention in many research fields. With the advent of low-cost digital signal processors, wireless sensor networks have begun to emerge in many applications. The majority of military applications, including our particular choice of tracking of personnel through a field of distributed sensors, possess a common core of signal processing functions. Because such sensor networks will be laid down in an *ad hoc* configuration consisting of thousands of sensor nodes, accurate and scalable algorithms are critical. The algorithms and approaches that we are developing under ISP Phase II are thus expected to have wide applicability. For example, we are working closely with the Raytheon group that is demonstrating shooter localization under DARPA Information Exploitation Office (IXO) Networked Embedded System Technology (NEST) program. Self-localization is a significant computation challenge for NEST and an opportunity for technology transfer. .

Distributed Tracking Demonstration Preliminaries

We now provide a high level overview of the distributed tracking demonstrations. We divide the demonstration into a series of tasks that are critical for the demonstration and a set of tasks that would provide additional capability but are not critical to the accomplishing the demonstration. We refer to these latter tasks as “Extras.” As illustrated in Figure 1, the distributed tracking demonstration consists of three functional blocks: (i) self-localization of the motes; (ii) 1-Bit on-mote detector; and (iii) base station tracker. These three functional blocks are discussed in more detail below and flow into a schedule as shown in Table 1.

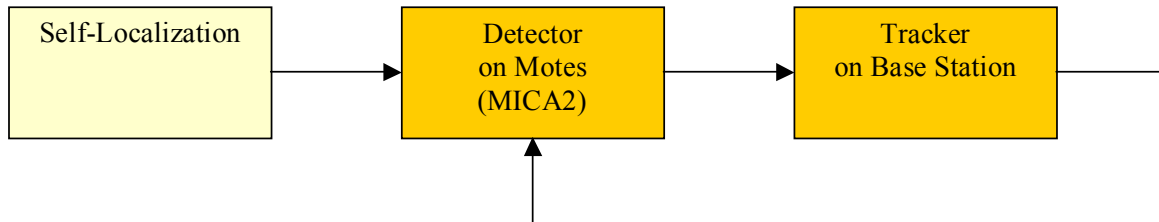


Figure 1: Demonstration Block Diagram

Self-Localization:

- Survey – If available self-localization algorithms do not produce accurate enough results, we should just localize motes by survey them.
- [Extra] Acoustic Ranging – VU algorithm currently gives reasonable results for inter-mote distance of 9 ft. For inter-mote distance higher than 9 ft, parameters need to be tweaked to reduce error in measured ranges.
- RIPS – The code needs to be installed onto MICA2’s. We may require permission to obtain UNCLASSIFIED code since it was developed under NEST program. Once installed, we need to make measurements behind M09 and evaluate results accuracy. Accuracy should be better than Acoustic Ranging Algorithm.
 - Drawbacks
 - [Extra] Current scheduling during data collection is too exhaustive and time consuming to make this a practical algorithm. For example, for only 16-mote network, data collection takes anywhere from 30 to 40 minutes. UniMelb wants to take this problem to improve scheduling by making only necessary measurements.
 - [Extra] Once measured data is collected, motes are localized using a Genetic Algorithm (GA). GA’s are known to be computationally intensive (thus, slow to converge to a solution), and they do not always converge. UM will investigate replacing the GA with a more reliable and faster algorithm such as the steepest descent.
- [Extra] RSSI – Determine the accuracy of this Received Signal Strength (RSS) algorithm from UM. Also, make sure UM can implement it on MICA2’s.

Detector (1-bit):

We next briefly discuss our detector development. The baseline demonstration will use a 1-Bit detector (target detected or not). This choice of detector implementation is driven by the network being so constrained in its communication capability. As shown in Figure

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

2, the detector is composed of four functional blocks: (i) Sensor Characterization; (ii) Signal Filtering; (iii) Energy Computation; and (iv) Threshold Calculation. For our scenario of tracking people through the network, sensor characterization consists of developing the acoustic signature of footsteps. Filtering is next performed to improve the Signal to Noise Ratio. A threshold is then set to produce the 1-Bit output of target detected or not. These four functional blocks are discussed in more detail below and again flow into a schedule as shown below.

Task #	Task Description	System Component	Performer	From	To	Duration (week)	Bonus/Mandatory
1	Interface Specification Document (Matlab/nesc)	ISD	All	2/20/2006	3/3/2006	1.57	M
	Motes Self-Localization (Acoustic Ranging)	Localization	Thom/Sal	-----	5/31/2006	2	B
	Motes Self-Localization (RIPS)	Localization	Bob/Sal	-----	5/31/2006	2	B
	Motes Self-Localization Data Collection Improvement (RIPS)	Localization	Craig/UniMelb	-----	5/31/2006	4	B
	Motes Self-Localization GA Replacement (RIPS)	Localization	Bob/Michigan	-----	5/31/2006	4	B
	Motes Self-Localization Implementation/Evaluation (RSSI)	Localization	Bob/Michigan	-----	5/31/2006	2	B
2	Sensor Characterization (Microphone) Acoustic Model/ROC's	Detector	Sal/ASU	2/13/2006	2/27/2006	2	M
	Sensor Characterization (Accelerometer) Vibration Model/ROC's	Detector	Sal/Thom	-----	4/17/2006	2	B
	Sensor Characterization (Magnetometer) Magnetic Model/ROC's	Detector	Sal/Thom	-----	4/17/2006	2	B
3	Filter Acoustic Footstep (Microphone)	Detector	Sal/ASU	2/27/2006	3/20/2006	3	M
	Filter Vibration Footstep (Accelerometer)	Detector	Sal/Thom	-----	4/17/2006	2	B
	Filter Magnetic Noise (Magnetometer)	Detector	Sal/Thom	-----	4/17/2006	2	B
4	Energy Computation (Microphone)	Detector	Sal/ASU	3/20/2006	3/27/2006	1	M
	Energy Computation (Accelerometer)	Detector	Sal/Thom	-----	4/17/2006	1	B
	Energy Computation (Magnetometer)	Detector	Sal/Thom	-----	4/17/2006	1	B
5	Threshold (Microphone)	Detector	Sal/ASU	3/27/2006	4/3/2006	1	M
	Threshold (Accelerometer)	Detector	Sal/Thom	-----	4/17/2006	1	B
	Threshold (Magnetometer)	Detector	Sal/Thom	-----	4/17/2006	1	B
6	Transmit 1-bit Detection	Detector	Sal/ASU	4/3/2006	4/17/2006	2	M
7	Tracker Single Target (Particle Filter)	Tracker	Sal/Thom/ASU	4/17/2006	5/31/2006	6.28	M
	Tracker Unknown Number of Targets (Particle Filter)	Tracker	Sal/ASU	2/1/2006	5/31/2006	17	B
8	Tracker (UniMelb)	Tracker	Craig/UniMelb	4/20/2006	5/31/2006	5.85	M
9	Detector/Tracker Integration	Integration	All	5/31/2006	6/30/2006	4.28	M
10	Motes Localization (Survey)	Localization	Bob/Thom/Sal	11/27/2006	12/1/2006	0.57	M
11	Full Dress Rehearsal Test	Testing	Bob/Thom/Sal	12/4/2006	12/15/2006	1.57	M

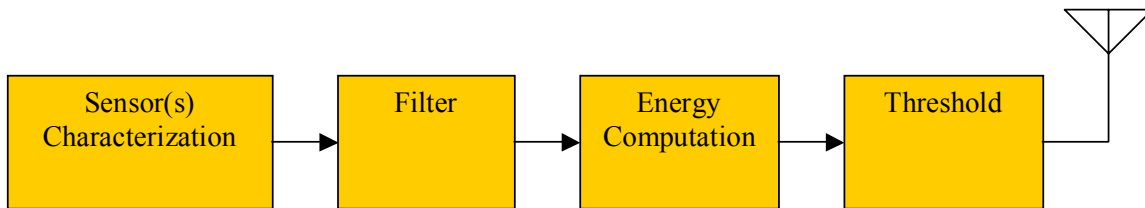


Figure 2: Detector Block Diagram

- Sensor(s) Characterization
 - Microphone (acoustic)
 - Person Walking – Determine ROC's to determine detector parameters, and motes network topology.
 - Accelerometer
 - [Extra] Person Walking – Determine if the sensor is capable of sensing vibration above noise floor on outdoor ground. If so, determine ROC's to determine detector parameters, and motes network topology.
 - Magnetometer
 - [Extra] Person Walking with Metal/Cell Phone – Determine if sensor can sense Metal or Cell Phones magnetic field. If so, determine ROC's to determine detector parameters, and motes network topology.
- Filter
 - Microphone (acoustic)
 - Person Walking – Develop Digital Filter similar to VU Acoustic Ranging Algorithm. The filter needs to be a Low-Pass. ASU will determine the frequency range of the filter.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

- Accelerometer
 - [Extra] Filtering can be ignored since vibrations are assumed to come only from people walking for the demo.
- Magnetometer
 - [Extra] Filtering can be ignored unless interference from Earth Magnetic Field or Magnetic Noise in the area affect detector.
- Energy Computation
 - It can be extracted or be similar to VU Acoustic Ranging Code
- Threshold
 - Determine threshold based on ROC's.
- Transmit Detection.
 - 1 – Target Detected
 - 0 – Target Not Detected

Tracker:

- Received Data – Receive detected/not detected data from each mote.
- Track – Track target using:
 - Particle Filter
 - Australian Tracker

Software:

- Matlab – Use Matlab to integrate Demo components and display tracker's graphics/results.
- NESC – Use to implement Localization and Detector algorithms on MICA2's.

Interface Specification Document:

In the next few weeks, we should have an Interface Specification Document describing the signals, variables, *etc.*, needed at the interface of each component of Figure 1. This document will describe, for example, the data and signals that detector needs to provide to the tracker.

Summary:

Figure 3 represents the final demonstration. It will have 40 to 100 sensors (S_1, \dots, S_n) detecting a target and possibly multiple targets. There will be one or more Base Stations depending on the number of available trackers. The Base Stations (trackers) will graphically show the target location using Matlab interface

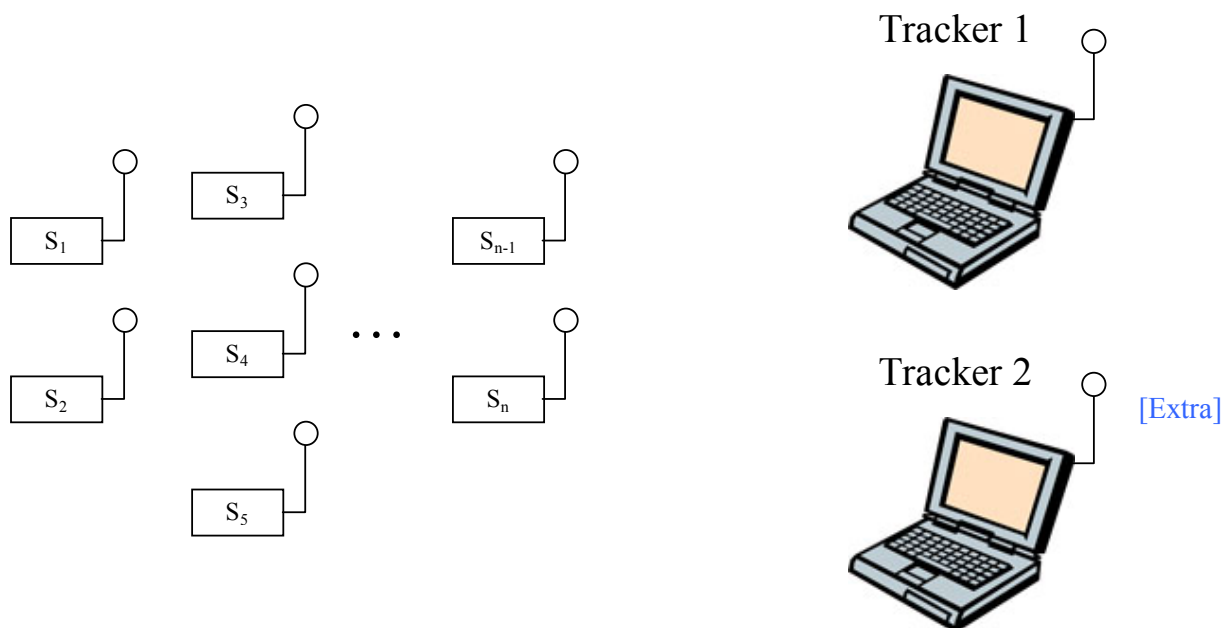


Figure 3: Demonstration Representation

The focus of Raytheon's technical efforts in support of the distributed tracking demonstration for the current PoP is self-localization.

The baseline self-localization algorithm relies on acoustic ranging and was developed by Vanderbilt University (VU). The concept of this algorithm is based on measuring the time of arrival (TOA) of the sound signal between the signal source (actuator) and the acoustic sensor. VU has also developed an approach that uses radio frequency instead of acoustic frequency and should significantly improved localization accuracy. Both algorithms rely on a genetic algorithm-based optimization approach which scales very poorly with the number of sensor nodes. We have evaluated the acoustic self-localization algorithm and found acceptable performance for an inter-mote spacing of ~3m. There are indications from the NEST program, that performance should be maintained out to ~10m and this is under investigation. As an alternative to the VU self-localization approach, we are investigating graph based algorithms motivated by concepts we are exploring for the processing of high dimensional data. Preliminary results are presented next and in the UM technical section.

2.A.1.b. Distributed Sensor Self-Localization

As described in previous quarterly reports for ISP phase II, self-localization is a key component of a wide variety of distributed wireless sensing applications, including perimeter monitoring, detection and tracking of targets, and shooter localization. Because such sensor networks will be laid down in an *ad hoc* configuration consisting of large numbers of sensor nodes, accurate and scalable localization algorithms are critical to the success of defense or homeland security applications.

The current generation of shooter localization algorithm is an acoustic ranging algorithm introduced by workers at Vanderbilt University (VU). The concept of this algorithm is based on measuring the time of arrival (TOA) of the sound signal between the signal source (actuator) and the acoustic sensor. The acoustic ranging algorithm has

demonstrated localization accuracy sufficient for a proof-of-principle, and VU is developing an approach that should significantly improve localization accuracy. This new approach uses radio frequency instead of acoustic frequency for the ranging algorithm [Maroti, et al. 2005], and should provide more accurate localization with larger networks since radio waves propagate further than acoustic waves. However, both the baseline and improved VU self-localization algorithms rely on an optimization approach based on genetic algorithms, which scale very poorly with the number of sensor nodes. Genetic algorithms (GA) suffer from slow convergence, as well as being awkward to implement in a distributed manner. This approach was selected only as a “first cut” by the research group at VU, and as currently implemented the computation is carried out on a base station (laptop computer), not distributed among the sensor nodes. Thus, improvement in the self-localization implementation is clearly needed, and we are working towards fulfillment of this goal.

In order to replace the GA approach to localization using acoustic TOA measurements, we have chosen to implement a version of “distributed, weighted multi-dimensional scaling” (dwMDS) algorithm, introduced by workers at the University of Michigan [Costa et al 2006]. “Classical” multi-dimensional scaling, or MDS, is a well-known algorithm, which has been popular for many years, and is a method of assigning a coordinate system to a group of objects for which we have some measure of “similarity” between pairs of objects. In the self-localization application, the similarity measurements comprise distances between pairs of nodes, and the MDS algorithm is well-suited to computing the coordinates of the nodes, up to an isometry, that is a rigid translation and/or rotation and/or reflection. To remove the ambiguity, it is necessary to have a few nodes for which the coordinates are known from outside measurements. These nodes are commonly called “anchor nodes,” and may be equipped with GPS, for example. As they are expensive, it is desirable to employ only a small number of them. The smallest number theoretically possible for removing the ambiguity in three dimensions is four; the number is $d + 1$ if the dimension of the underlying coordinate space is d . Since MDS can compute the coordinates for each node without knowing the exact locations of any of the nodes, the information from the anchor nodes can be utilized after this part of the computation has been completed, in the form of a simple matrix multiplication and/or vector addition applied to each coordinate vector. However, it may be more efficient to incorporate this prior information during computation of the coordinates, and classical MDS is not particularly well-suited to incorporating this additional information. Furthermore, classical MDS requires knowledge of all pair-wise distances between nodes, and is not applicable if only some of the pair-wise distances are known, for example between nearest neighbors. In addition, due to the presence of noise, it may be advantageous to weight the distance measurements differently, to reflect their different levels of reliability. This classical MDS also cannot do. The dwMDS algorithm overcomes all of these deficiencies, and we are currently implementing this algorithm for use in the distributed sensor network using either TOA or received signal strength (RSS) measurements.

In addition to TOA or RSS measurements, we also have available a form of radio interferometer measurements, due to a method introduced by the researchers at VU [Maroti, et al. 2005]. None of the usual approaches to self-localization are suitable for using these measurements directly, as they are not pair-wise distances, but a combination

of four pair-wise distances. The measurement is made by using two nodes as transmitters and two nodes as receivers. The transmitters emit radio signals at slightly different frequencies, and the phase difference of the beat frequency is measured at the two receivers. After a bit of post processing to estimate the number of wavelengths required to reach the receivers, what we obtain is

$$R_{ijkl} = d_{il} - d_{ik} + d_{jk} - d_{jl} , \quad (1)$$

where (i, j) are transmitters, (k, l) are receivers, and d_{ij} is the pair-wise distance between nodes i and j . We refer to the radio interferometer measurements on the left-hand side of (1) as RIMs. Since the beginning of the year we have been studying the form of the RIMs, to decide what can be done to replace the genetic algorithm approach here. That it is desirable to do so is due to the fact that the radio measurements are more accurate than RSS, and are more reliable over a longer distance than acoustic TOA measurements, if these are used for localization (acoustic TOA would still be used for shooter localization or tracking algorithms, the localization part is a separate issue).

The principle difficulty with RIMs, ignoring any potential issues with noise, etc., is that the number of independent measurements is only $N(N-3)/2$, where N is the number of nodes. This has been proved in [Meertens 2005]. The number of pair-wise distances on the right-hand side of (1) is $N(N-1)/2$, hence there is a null space of dimension N . This means that any attempt to solve the system of equations consisting of all measurements (1) will fail, since the right- and left-hand sides of (1) can be made to agree, even with estimates of the positions of the nodes which are far from the true positions. This is due to the fact that the system is under-determined. To overcome this difficulty it will be necessary to design an algorithm which utilizes additional information, that is, in addition to the RIMs measurements. N. Patwari at University of Michigan has suggested a combination of RIMs and RSS measurements [Patwari 2006], and it seems to us that this is a very good idea. Such an approach is currently under investigation and we hope to report favorable results soon.

2.A.1.c. Evaluation of Distribution-Free Divergence Measurements

Introduction

In working with reconfigurable or agile sensors capable of producing different feature sets, it is useful to quantify which feature sets provide the best opportunity to discriminate and classify targets of interest. When there are more than approximately 5 features, the feature set comparisons take on a “high-dimensional” nature. Limited sample support in high-dimensional spaces leads to the well known “curse of dimensionality.” Distribution-based comparisons of feature-set efficacy are prone to error, whether the distribution estimation is based on density-building using kernels or fitting parameters in a predefined model. There are some methods available for distribution-free, approximate measurement of feature-set divergence (or separability). In the two-class case, the measurement of feature-set efficacy can be recast in terms of the general multivariate two-sample problem. The underlying assumption is that independent of any classifier, feature sets that exhibit more divergence (or separability) should in general be of greater utility than feature sets that exhibit less divergence (or separability).

Assume we are given two independent \mathcal{R}^d -valued random vectors $\mathbf{X}_1, \dots, \mathbf{X}_{n_1}$ and $\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}$, with the distribution of \mathbf{X}_i having the unknown pdf $f(\mathbf{x})$ and the distribution of \mathbf{Y}_i having the unknown pdf $g(\mathbf{x})$. The multivariate two-sample problem is to test the hypothesis: $H_0 : f = g$ versus the general alternative. For the multivariate case, several approximate distribution-free methods are known, including:

- 1) α Jensen Rényi Divergence [Hero et al. 2002]
- 2) Henze-Penrose Divergence [Henze & Penrose 1999], [Friedman & Rafsky 1979]
- 3) Henze-Schilling Statistic [Henze 1988], [Schilling 1986]
- 4) Euclidean Interpoint Distance Test [Baringhaus & Franz 2004], [Székely & Rizzo 2004]; Statistical Energy Test [Aslan & Zech 2005]
- 5) α Geometric-Arithmetic Mean Test [Neemuchwala & Hero 2005]
- 6) α Mutual Information Test [Neemuchwala & Hero 2005]
- 7) Dwyer-Squire Test [Dwyer & Squire 1993]
- 8) Change Point Test [Ferber 2000].
- 9) Location-Scale Test [Rousson 2001]
- 10) Incomplete Multivariate Observations Test [Wei & Lachin 1984]

The first three techniques are considered here; the remaining will be evaluated in future work.

The two feature sets being compared are composed of unit-variance normal distributions, with differing means, dimensionality and sample size (assumed throughout to be the same for the two feature sets, $n_1 = n_2$). We are testing against location alternatives; in future work we will test against scale alternatives, as well as use different types of random distributions. In Figure 4 we show some examples of $d=2$ feature sets composed of unit-variance normal distributions with mean separations varying from 0 units to 8 units, for a fixed sample size $n_1 = n_2 = 500$. In Figure 5 we show some examples of $d=2$ feature sets composed of unit-variance normal distributions with a fixed mean separation of 3 units, for sample sizes of $n_1 = n_2 = 100, 500, 4000$. We repeat each experiment ten times, and report the mean and standard deviation of the divergence estimate.

α Jensen Rényi Divergence

The α Jensen Rényi Divergence uses the Rényi entropies of the individual feature sets and compares them with the Rényi entropy of the combined feature set:

$$\Delta H_\alpha(f, g) \equiv H_\alpha(f + g) - 0.5(H_\alpha(f) + H_\alpha(g)); \quad H_\alpha(f) \equiv \frac{1}{1-\alpha} \ln \int_z f^\alpha(z) dz.$$

The individual Rényi entropies are estimated, in an asymptotic manner, using the minimal spanning tree (MST) of a graph having power-weighted edges connecting the d -dimensional sample nodes. There is one free parameter, $0 < \alpha < 1$. This parameter is related to the power weighting exponent, γ , and the dimensionality d in the following manner: $\gamma = d(1-\alpha)$. Values of α close to 0 produce large exponents. This amplifies the effects of the longest edge length, thus tending to emphasize tail differences between the two distributions. Values of α close to 1 produce small exponents. This provides similar

treatment to different values of edge lengths, thus tending to emphasize central differences between the two distributions (where many more edges are present).

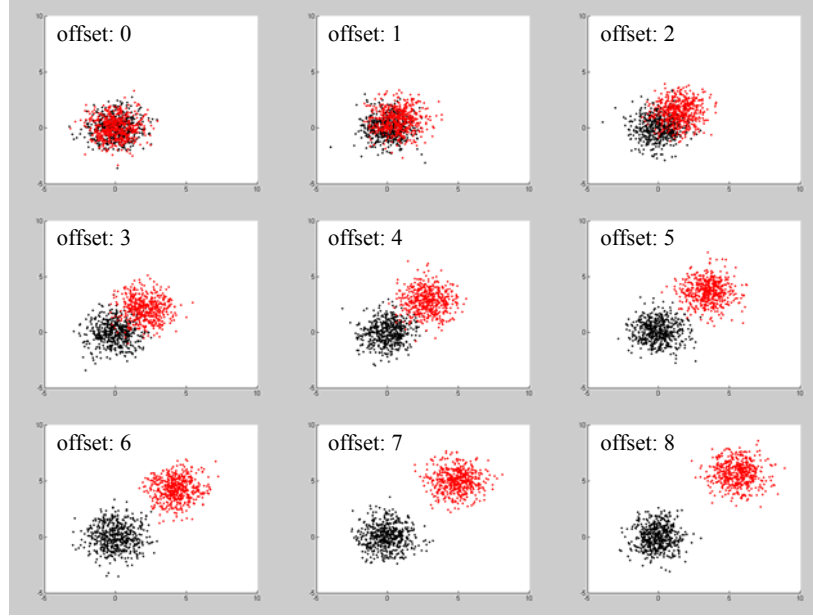


Figure 4: Examples of two-class feature set distributions for $d=2$, $n_1=n_2=500$. The distributions are drawn from unit-variance normals, with means offsets varying from 0 units to 8 units as labeled.

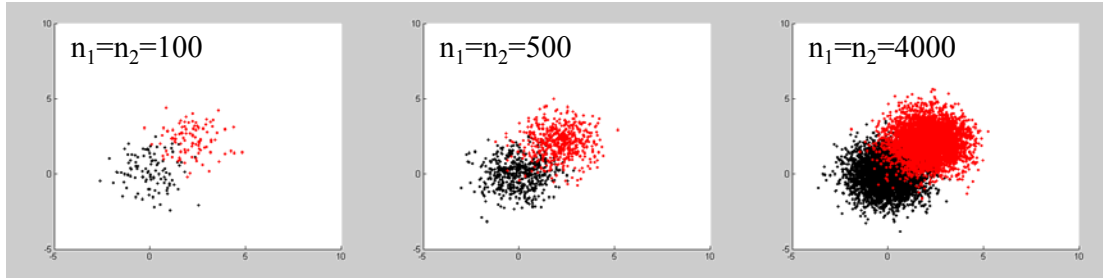


Figure 5: Examples of two-class feature set distributions for $d=2$, $n_1=n_2=100, 500, 4000$. The distributions are drawn from unit-variance normals having a mean offset of 3 units.

We first examine the behavior of ΔH_α for fixed value $\alpha=0.5$, for two unit-variance normals whose means are separated by various amounts, and for various sample sizes. Each scenario is repeated with 10 realizations of the sample sets, and the mean and standard deviations of ΔH_α are shown in Figure 6 for $d=2$ and in Figure 7 for $d=20$. It appears that for $d=2$, the asymptotic results are approached with a few thousand samples in each feature set, but for sample sizes < 500 , there is substantial variability and overlap in divergence values for different amounts of feature-set separation, and the measures have not reached the large-sample-size asymptotic values. For $d=20$, it appears that even with several thousand samples in each feature set, the asymptotic values have not been achieved for large separation values, and for sample sizes < 500 there is much inconsistency (including a negative bias for low-separation estimates of $\Delta H_\alpha \geq 0$).

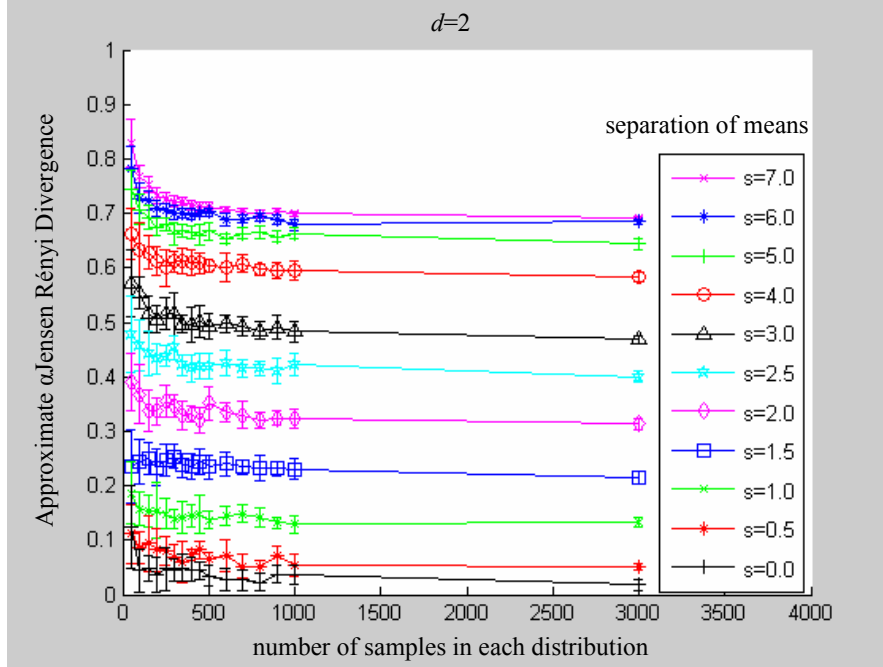


Figure 6: Approximate α Jensen Rényi Divergence values for two unit-normal distributions ($d=2$) with different mean separations, as a function of sample size, for $\alpha=0.5$ (mean and std plotted for 10 realizations).

We next fix $n_1 = n_2 = 2000$ and vary α for $d=2$ and $d=20$, again using unit-variance normal distributions with differing amounts of mean separation. Results are shown in Figure 8 for $d=2$ and Figure 9 for $d=20$. Note that for low separation values, there is not much sensitivity to either α or d . For large separation values, there is a dependence on α that becomes stronger as d increases. This is to be expected, as the edge lengths in the MST are power-weighted by the exponent $\gamma = d(1-\alpha)$. For larger values of d and smaller values of α , the power-weighting exponent is larger, emphasizing the maximum edge length. For separated distributions, this corresponds to the edge-to-edge distance between the two clusters.

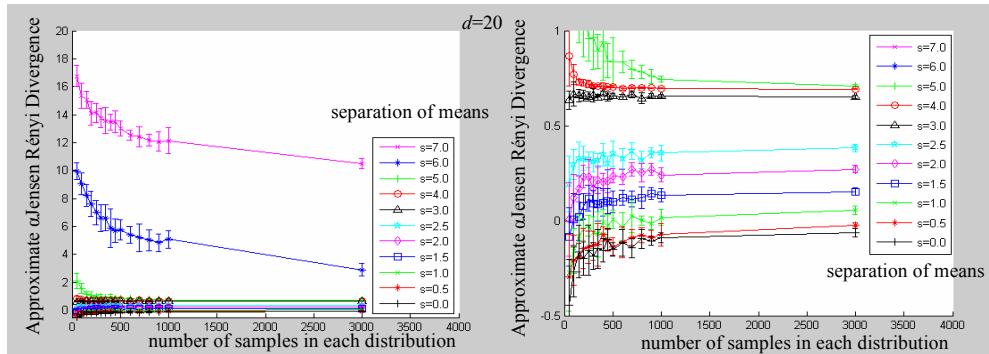


Figure 7: Approximate α Jensen Rényi Divergence values for two unit-normal distributions ($d=20$) with different mean separations, as a function of sample size, for $\alpha=0.5$ (mean and std plotted for 10 realizations).

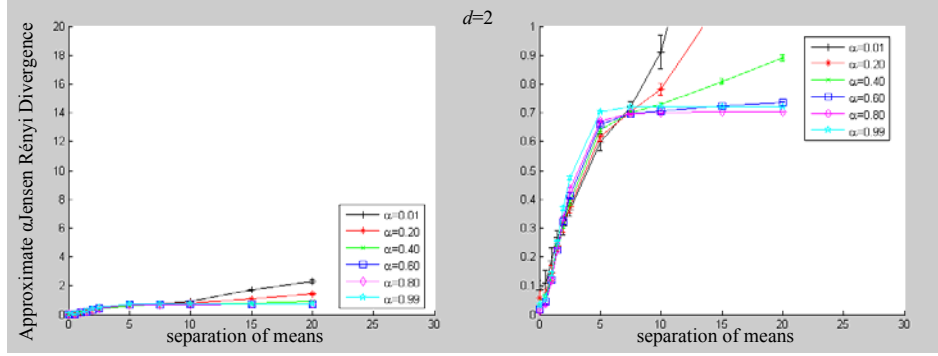


Figure 8: Approximate α Jensen Rényi Divergence values for two unit-normal distributions ($d=2$) with different α values, as a function of mean separation, for sample size = 2000 samples in each distribution (mean and std plotted for 10 realizations).

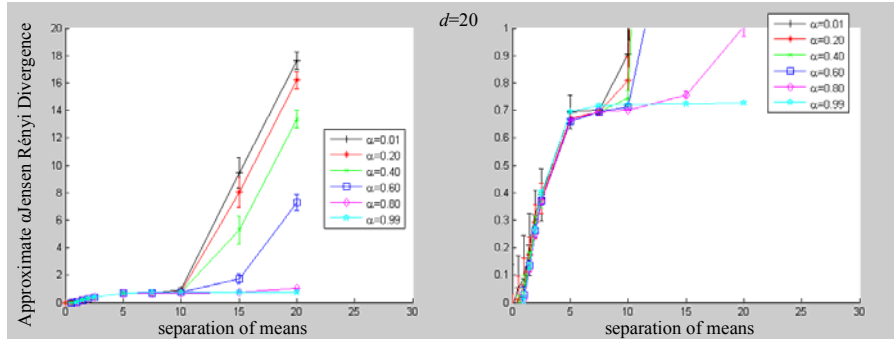


Figure 9: Approximate α Jensen Rényi Divergence values for two unit-normal distributions ($d=20$) with different α values, as a function of mean separation, for sample size = 2000 samples in each distribution (mean and std plotted for 10 realizations).

Henze-Penrose Divergence

Friedman and Rafsky [Friedman & Rafsky 1979] report a distribution-free generalization of the Wald-Wolfowitz runs statistic. The feature sets $\mathbf{X}_1, \dots, \mathbf{X}_{n_1}$ and $\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}$ are pooled, and the Minimal Spanning Tree (MST) of the resulting data set is formed. In the simplest form of the test, count the number R of MST edges that join nodes from different samples. In a further refinement, Friedman & Rafsky allow usage of multiple orthogonal MSTs. In this case, count the number S of edges in the multiple orthogonal MSTs that join nodes from different samples. A prescription is given for estimating the expected mean and variance of R or S under the null hypothesis $H_0: f = g$, allowing the estimation of the z-score of R or S as a measure of the violation of the null hypothesis. Henze and Penrose [Henze & Penrose 1999] show that asymptotically in the sample size, the statistic R can be related to $\delta(f, g, p)$, a particular measure of distributional divergence (where $p = \frac{n_1}{n_1 + n_2}$):

$$\delta(f, g, p) \equiv \int \frac{p^2 f^2(x) + (1-p)^2 g^2(x)}{p f(x) + (1-p) g(x)} dx; \quad \{p^2 + (1-p)^2\} \leq \delta(f, g, p) \leq 1.$$

In the case where $n_1 = n_2 = n$, we have $p=0.5$ so $0.5 \leq \delta \leq 1$, and

$$\delta(f, g, p) = \lim_{n \rightarrow \infty} \left(1 - \frac{R}{2n} \right).$$

Henze and Penrose do not comment on a connection between $\delta(f, g, p)$ and S . It appears plausible (we will investigate in future work) that one can use S for estimating $\delta(f, g, p)$, for ‘small’ values of k = number of orthogonal MSTs used, with the modification

$$\delta(f, g, p) \approx \lim_{n \rightarrow \infty} \left(1 - \frac{S}{2nk} \right).$$

We use the same distributions used for evaluating the α Jensen Rényi Divergence. In Figure 10, we plot approximate Henze-Penrose Divergence results for $d=2$. Note that as with the approximate α Jensen Rényi Divergence values for $d=2$, there is some overlap in values at the lowest separations. However the mean values for the approximate Henze-Penrose Divergence measures at low sample number are consistent with the asymptotic values, whereas the low-sample-number α Jensen Rényi Divergence values show bias with respect to the asymptotic values. Also, in contrast with the α Jensen Rényi Divergence, the Henze-Penrose Divergence does not continue to increase once the two feature set clusters are separated. In Figure 10, we plot approximate Henze-Penrose Divergence results for $d=20$. For sample size < 500 , there is some bias relative to the asymptotic values, but both the bias and the variance appear to be less than that for the α Jensen Rényi Divergence values (Figure 7).

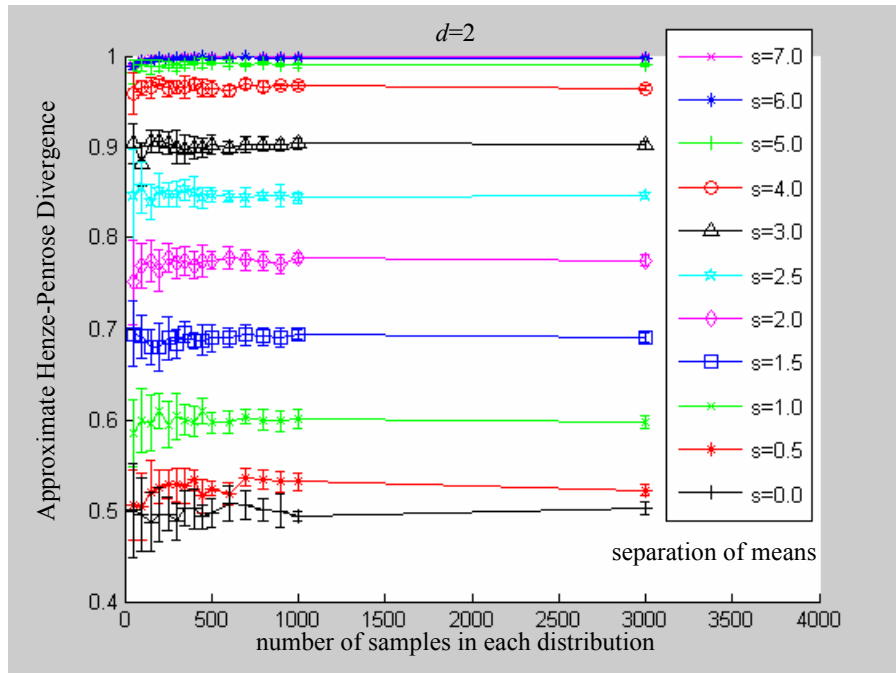


Figure 10: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=2$) with different mean separations, as a function of sample size, for $k=1$ (mean and std plotted for 10 realizations).

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

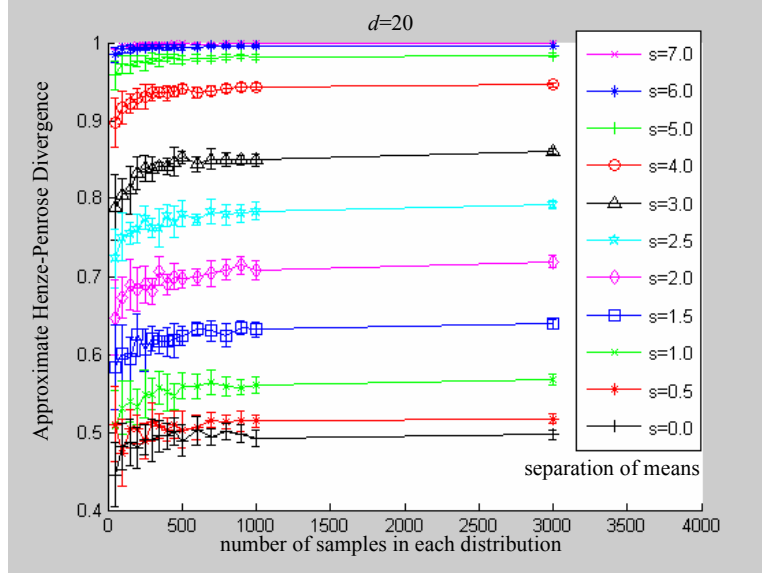


Figure 11: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=20$) with different mean separations, as a function of sample size, for $k=1$ (mean and std plotted for 10 realizations).

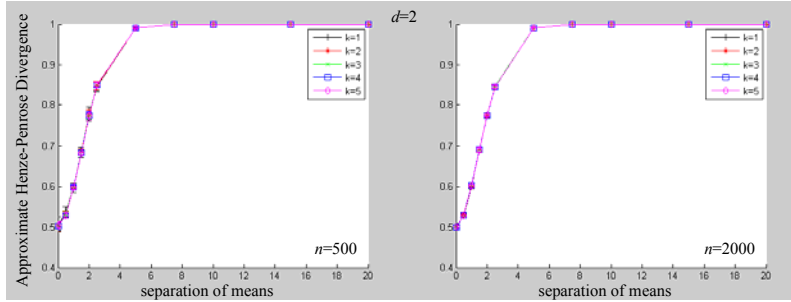


Figure 12: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=2$) with sample sizes of $n_1=n_2 \in \{500, 2000\}$, as a function of mean separation, for $k \in \{1, 2, 3, 4, 5\}$ (mean and std plotted for 10 realizations).

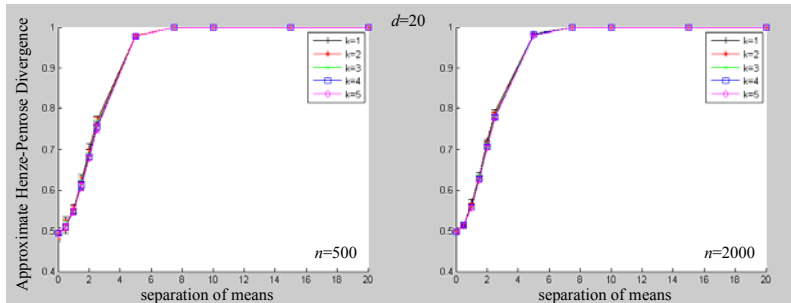


Figure 13: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=20$) with sample sizes of $n_1=n_2 \in \{500, 2000\}$, as a function of mean separation, for $k \in \{1, 2, 3, 4, 5\}$ (mean and std plotted for 10 realizations).

We next use $n_1 = n_2 \in \{500, 2000\}$ and vary $k \in \{1, 2, 3, 4, 5\}$ for $d=2$ and $d=20$, again using unit-variance normal distributions with differing amounts of mean separation.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

Results are shown in Figure 12 for $d=2$ and Figure 13 for $d=20$. For $d=2$, the choice of $n_1 = n_2 \in \{500, 2000\}$ or $k \in \{1, 2, 3, 4, 5\}$ produces no meaningful difference in the mean values, as shown in Table 1. The standard deviations are smaller in the $n_1 = n_2 = 2000$ case compared with the $n_1 = n_2 = 500$ case, and the standard deviations decrease as k increases, as shown in Table 2. For $d=20$, the choice of $n_1 = n_2 \in \{500, 2000\}$ or $k \in \{1, 2, 3, 4, 5\}$ does produce a small but measurable change in the mean values, as shown in Table 3. The standard deviations are smaller in the $n_1 = n_2 = 2000$ case compared with the $n_1 = n_2 = 500$ case, and the standard deviations decrease as k increases, as shown in Table 4. Also, for the lower separation values (where the two distributions are actually overlapped), the $d=20$ approximate Henze-Penrose Divergence values are lower than the $d=2$ approximate Henze-Penrose Divergence. This is expected, as the unit-variance distributions are more spread out in higher dimensions (as dimensionality increases, less mass is contained within one standard deviation). Since the two distributions are more spread out in $d=20$ than in $d=2$, for a given mean separation there is more overlap and thus the approximate Henze-Penrose Divergence is less.

$d=2$	$n=500, k=1$	$n=500, k=2$	$n=500, k=3$	$n=500, k=4$	$n=500, k=5$	$n=2000, k=1$	$n=2000, k=2$	$n=2000, k=3$	$n=2000, k=4$	$n=2000, k=5$
$s=0.0$	0.507	0.502	0.503	0.503	0.504	0.502	0.500	0.500	0.499	0.500
$s=0.5$	0.537	0.532	0.532	0.529	0.529	0.527	0.530	0.529	0.528	0.528
$s=1.0$	0.596	0.598	0.598	0.597	0.597	0.600	0.601	0.601	0.602	0.601
$s=1.5$	0.683	0.685	0.684	0.683	0.681	0.691	0.691	0.690	0.690	0.689
$s=2.0$	0.778	0.777	0.776	0.775	0.774	0.774	0.775	0.774	0.775	0.776
$s=2.5$	0.846	0.848	0.850	0.850	0.850	0.847	0.847	0.846	0.845	0.845
$s=5.0$	0.991	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991	0.991
$s=7.5$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
$s=10.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
$s=15.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
$s=20.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000

Table 1: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=2$) with sample sizes of $n_1=n_2 \in \{500, 2000\}$, as a function of mean separation, for $k=\{1, 2, 3, 4, 5\}$, mean for 10 realizations.

$d=2$	$n=500, k=1$	$n=500, k=2$	$n=500, k=3$	$n=500, k=4$	$n=500, k=5$	$n=2000, k=1$	$n=2000, k=2$	$n=2000, k=3$	$n=2000, k=4$	$n=2000, k=5$
$s=0.0$	0.019	0.012	0.010	0.010	0.009	0.007	0.004	0.004	0.003	0.003
$s=0.5$	0.014	0.007	0.005	0.004	0.006	0.006	0.005	0.002	0.003	0.003
$s=1.0$	0.012	0.006	0.006	0.005	0.005	0.006	0.005	0.005	0.004	0.004
$s=1.5$	0.012	0.008	0.006	0.005	0.003	0.006	0.007	0.006	0.007	0.006
$s=2.0$	0.017	0.011	0.010	0.011	0.011	0.007	0.005	0.004	0.004	0.003
$s=2.5$	0.011	0.010	0.007	0.007	0.006	0.005	0.005	0.005	0.004	0.004
$s=5.0$	0.004	0.003	0.003	0.003	0.003	0.001	0.001	0.001	0.001	0.001
$s=7.5$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=10.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=15.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=20.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 2: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=2$) with sample sizes of $n_1=n_2 \in \{500, 2000\}$, as a function of mean separation, for $k=\{1, 2, 3, 4, 5\}$, std. dev. for 10 realizations.

$d=20$	$n=500$	$n=500$	$n=500$	$n=500$	$n=500$	$n=2000$	$n=2000$	$n=2000$	$n=2000$	$n=2000$
--------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
$s=0.0$	0.491	0.490	0.493	0.494	0.496	0.499	0.499	0.498	0.498	0.499
$s=0.5$	0.514	0.514	0.513	0.511	0.510	0.514	0.515	0.515	0.515	0.513
$s=1.0$	0.554	0.552	0.548	0.548	0.547	0.566	0.559	0.557	0.556	0.556
$s=1.5$	0.620	0.619	0.614	0.611	0.609	0.635	0.632	0.630	0.627	0.625
$s=2.0$	0.699	0.688	0.684	0.681	0.678	0.720	0.714	0.709	0.707	0.704
$s=2.5$	0.771	0.763	0.759	0.756	0.752	0.792	0.787	0.784	0.781	0.779
$s=5.0$	0.980	0.980	0.978	0.978	0.977	0.983	0.982	0.982	0.981	0.981
$s=7.5$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	0.999	0.999	0.999
$s=10.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
$s=15.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
$s=20.0$	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000

Table 3: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=20$) with sample sizes of $n_1=n_2 \in \{500,2000\}$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, mean for 10 realizations.

$d=20$	$n=500, k=1$	$n=500, k=2$	$n=500, k=3$	$n=500, k=4$	$n=500, k=5$	$n=2000, k=1$	$n=2000, k=2$	$n=2000, k=3$	$n=2000, k=4$	$n=2000, k=5$
$s=0.0$	0.017	0.010	0.006	0.005	0.004	0.009	0.006	0.005	0.004	0.003
$s=0.5$	0.020	0.014	0.011	0.008	0.009	0.008	0.008	0.006	0.004	0.003
$s=1.0$	0.011	0.010	0.008	0.008	0.008	0.010	0.008	0.007	0.007	0.006
$s=1.5$	0.018	0.014	0.012	0.010	0.009	0.007	0.005	0.005	0.005	0.005
$s=2.0$	0.015	0.013	0.008	0.008	0.006	0.004	0.005	0.006	0.006	0.006
$s=2.5$	0.011	0.014	0.012	0.011	0.010	0.005	0.004	0.003	0.004	0.004
$s=5.0$	0.004	0.003	0.004	0.004	0.004	0.001	0.001	0.002	0.002	0.001
$s=7.5$	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=10.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=15.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$s=20.0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4: Approximate Henze-Penrose Divergence values for two unit-normal distributions ($d=20$) with sample sizes of $n_1=n_2 \in \{500,2000\}$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, std. dev. for 10 realizations.

Henze-Schilling Test

[Henze 1988] extends previous work [Schilling1986] to present a distribution-free multivariate two-sample test. The Henze-Schilling test is based on $T(p,k)$, the cumulative number of k -nearest-neighbor type coincidences measured over the pooled sample set consisting of $p = n_1 + n_2$ samples. For a fixed but arbitrary norm $\|\cdot\|$ on \mathbb{R}^d and given the pooled sample set $Z_{1,\dots,p} = \{X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}\}$, define the r -th nearest neighbor to Z_i (denoted by $N_r(Z_i)$) as that point Z_j satisfying $\|Z_m - Z_i\| < \|Z_j - Z_i\|$ for exactly $r-1$ values of m , $1 \leq m \leq p$; $m \neq i, j$. Then define the indicator variable $I(i,r)$ as follows: $I(i,r) = 1$, if Z_i and $N_r(Z_i)$ belong to the same parent sample set; $I(i,r) = 0$ otherwise. The test statistic $T(n,k)$ is then defined as:

$$T(n,k) \equiv \sum_{i=1}^p \sum_{r=1}^k I(i,r).$$

Under the null hypothesis $H_0 : f = g$, it is shown that asymptotically $T(p,k)$ is normally distributed, and the expected mean and variance of $T(p,k)$ under the null hypothesis are given. One can then form the z-score of $T(p,k)$ and reject H_0 at approximate level α :

If $\frac{T - E(T|H_0)}{\text{std}(T|H_0)} \geq \Phi^{-1}(1 - \alpha)$, then reject H_0 .

The Henze-Schilling Statistic, for equal sample sizes $n_1 = n_2 = n$, is expressed as $\frac{T(p, k)}{2nk}$.

We use the same distributions as before. In Figure 14 and Figure 15, we plot the Henze-Schilling Statistic for $d=2$ and $d=20$, respectively. However due to run-time considerations we stop at $n_1 = n_2 = 800$ for $d=2$ and $n_1 = n_2 = 1000$ for $d=20$. In Figure 16 we plot the Henze-Schilling Statistic for $d=2$ and $d=20$, for fixed sample size $n_1 = n_2 = 500$ and for $k \in \{1, 2, 3, 4, 5\}$.

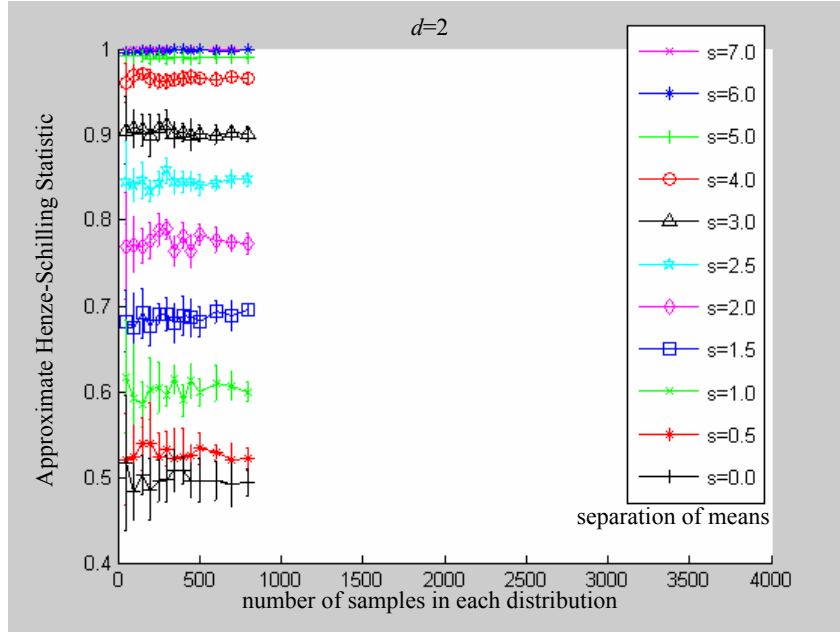


Figure 14: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=2$) with different mean separations, as a function of sample size, for $k=1$ (mean and std plotted for 10 realizations).

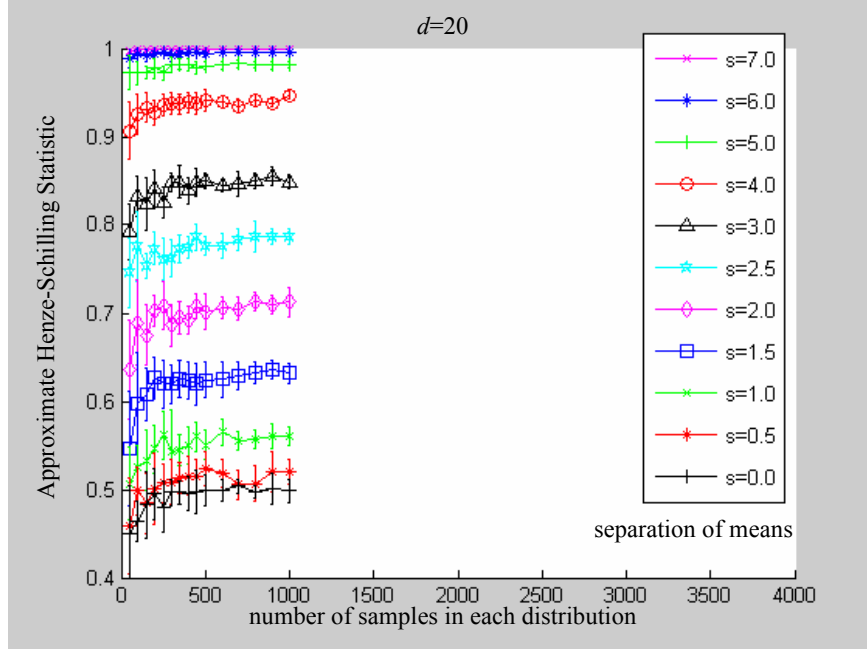


Figure 15: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=20$) with different mean separations, as a function of sample size, for $k=1$ (mean and std plotted for 10 realizations).

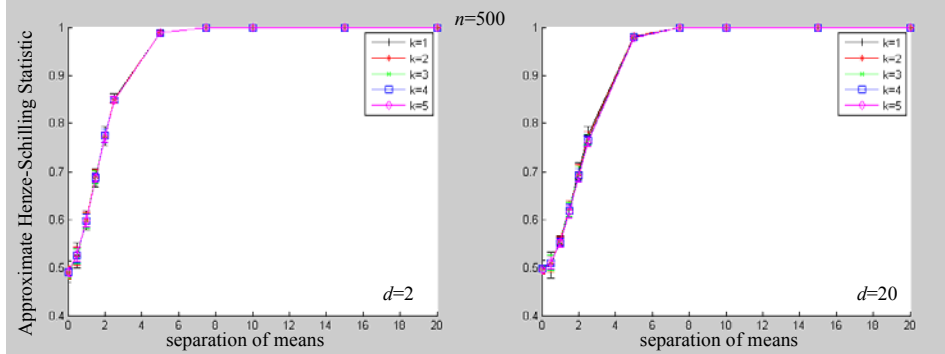


Figure 16: Approximate Henze-Schilling Statistic values for two unit-normal distributions with dimensionalities $d=2$ and $d=20$, for sample size $n_1=n_2=500$ as a function of mean separation, for $k \in \{1, 2, 3, 4, 5\}$ (mean and std plotted for 10 realizations).

Note that the Henze-Schilling Statistic is remarkably similar in value to the Henze-Penrose Divergence (compare Figure 14 and Figure 15 to Figure 10 and Figure 11, and compare Figure 16 to the left side plots in Figure 12 and Figure 13). This is to be expected. The Henze-Penrose Divergence makes use of the Friedman-Rafsky statistic S , which is a measure of how many edges in k orthogonal MSTs connect nodes from different parent distributions. The Henze-Schilling Statistic T is a measure of how many edges in the cumulative k -nearest-neighbor graph connect nodes from the same parent distribution. In both cases, the total number of edges in the fully-connected graphs is $(2n-1)k$, where $2n = n_1 + n_2$ in the case $n_1 = n_2 = n$ is the total number of nodes and k is the number of orthogonal graphs included, whether MSTs or k NNs. The first MST is

analogous to the 1-nn graph – nodes are connected to their nearest neighbor. Increasing k to $k=2$ adds in the second MST which is analogous to adding in the 2-nn graph, and so on. Normalizing both S and T by $2nk$, which for large n is practically equal to the total number of edges $(2n-1)k$, results in $S/2nk$ and $T/2nk$. The former is practically equal to the fraction of edges in k orthogonal MSTs that connect nodes from different distributions, and the latter is practically equal to the fraction of edges in cumulative k NNs that connect nodes from the same distribution. Subtracting the former from 1 gives, practically, the fraction of edges in k orthogonal MSTs that connect nodes from the same distribution, and of course $1 - S/2nk$ is simply the approximate Henze-Penrose Divergence.

Thus, to the extent that cumulative orthogonal MSTs use the same edges as cumulative k NN graphs, one would expect the Henze-penrose Divergence and the Henze-Schilling Statistic to have similar values. Friedman and Rafsky [Friedman & Rafsky 1979] allude to this correspondence: “Results were derived ... for the mean and variance of a runs statistic based on the MST. However, the derivations do not require that the set of edges considered form an MST or even a tree. The results are valid for any graph with exactly $N-1$ edges, and moreover, for any graph containing the N points... To have reasonable power against general alternatives, it is necessary that the edges generally link points that are close in the observation space. As pointed out, this motivated our choice of the MST. The graph that links every point to its nearest neighbor(s) is another possibility.” In Figure 17 we plot the approximate Henze-Schilling Statistic against the approximate Henze-Penrose Divergence (mean values from ten realizations of the random distributions), for fixed sample size $n_1 = n_2 = 500$ and for $k \in \{1, 2, 3, 4, 5\}$. The two divergence metrics are clearly highly correlated; in fact in all ten cases ($d=2$ and $d=20$, for $k \in \{1, 2, 3, 4, 5\}$), the approximate Henze-Penrose Divergence and approximate Henze-Schilling Statistic have a correlation coefficient $\rho > 0.99$. For ease of comparison of behavior of the means and standard deviations (from ten realizations of the random distributions) of the approximate Henze-Penrose Divergence and approximate Henze-Schilling Statistic, we tabulate the mean and standard deviation for the approximate Henze-Schilling Statistic in Tables 5–8. These values can be compared with those of the approximate Henze-Penrose Divergence, given in Tables 1–4.

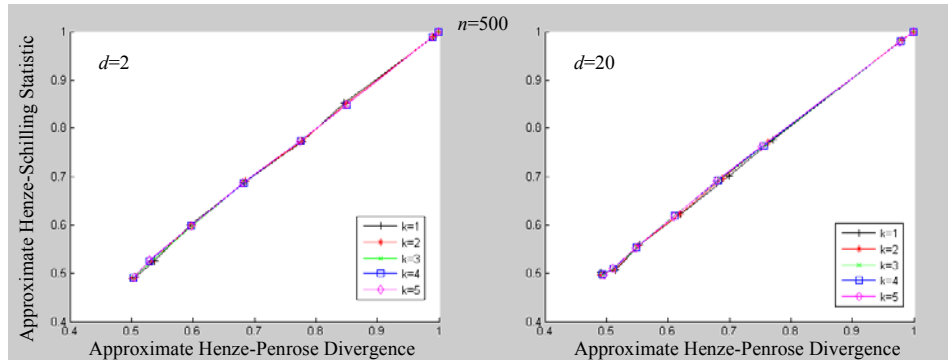


Figure 17: Approximate Henze-Schilling Statistic values plotted against approximate Henze-Penrose Divergence values, for two unit-normal distributions with dimensionalities $d=2$ and $d=20$, for sample size $n_1=n_2=500$ as a function of mean separation, for $k \in \{1, 2, 3, 4, 5\}$ (mean values from 10 realizations).

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

$d=2$	$n=500,$ $k=1$	$n=500,$ $k=2$	$n=500,$ $k=3$	$n=500,$ $k=4$	$n=500,$ $k=5$
$s=0.0$	0.492	0.489	0.490	0.491	0.491
$s=0.5$	0.526	0.524	0.524	0.525	0.527
$s=1.0$	0.598	0.598	0.597	0.597	0.597
$s=1.5$	0.687	0.690	0.689	0.686	0.687
$s=2.0$	0.773	0.773	0.774	0.774	0.773
$s=2.5$	0.852	0.850	0.849	0.849	0.849
$s=5.0$	0.989	0.989	0.989	0.988	0.988
$s=7.5$	1.000	1.000	1.000	1.000	1.000
$s=10.0$	1.000	1.000	1.000	1.000	1.000
$s=15.0$	1.000	1.000	1.000	1.000	1.000
$s=20.0$	1.000	1.000	1.000	1.000	1.000

Table 5: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=2$) with sample size of $n_1=n_2=500$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, mean for 10 realizations.

$d=2$	$n=500,$ $k=1$	$n=500,$ $k=2$	$n=500,$ $k=3$	$n=500,$ $k=4$	$n=500,$ $k=5$
$s=0.0$	0.022	0.013	0.012	0.009	0.009
$s=0.5$	0.026	0.019	0.016	0.016	0.015
$s=1.0$	0.021	0.016	0.016	0.013	0.012
$s=1.5$	0.020	0.014	0.014	0.010	0.009
$s=2.0$	0.020	0.015	0.015	0.014	0.011
$s=2.5$	0.010	0.007	0.008	0.008	0.008
$s=5.0$	0.004	0.004	0.004	0.005	0.004
$s=7.5$	0.000	0.001	0.000	0.000	0.000
$s=10.0$	0.000	0.000	0.000	0.000	0.000
$s=15.0$	0.000	0.000	0.000	0.000	0.000
$s=20.0$	0.000	0.000	0.000	0.000	0.000

Table 6: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=2$) with sample size of $n_1=n_2=500$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, std. dev. for 10 realizations.

$d=20$	$n=500,$ $k=1$	$n=500,$ $k=2$	$n=500,$ $k=3$	$n=500,$ $k=4$	$n=500,$ $k=5$
$s=0.0$	0.501	0.495	0.496	0.497	0.495
$s=0.5$	0.506	0.508	0.510	0.509	0.509
$s=1.0$	0.558	0.556	0.554	0.554	0.551
$s=1.5$	0.622	0.623	0.621	0.618	0.617
$s=2.0$	0.701	0.696	0.695	0.693	0.690
$s=2.5$	0.777	0.771	0.766	0.765	0.763
$s=5.0$	0.982	0.982	0.980	0.980	0.979
$s=7.5$	0.999	0.999	0.999	0.999	0.999
$s=10.0$	1.000	1.000	1.000	1.000	1.000
$s=15.0$	1.000	1.000	1.000	1.000	1.000
$s=20.0$	1.000	1.000	1.000	1.000	1.000

Table 7: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=20$) with sample size of $n_1=n_2=500$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, mean for 10 realizations.

$d=20$	$n=500,$	$n=500,$	$n=500,$	$n=500,$	$n=500,$
--------	----------	----------	----------	----------	----------

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
$s=0.0$	0.014	0.009	0.008	0.008	0.006
$s=0.5$	0.027	0.017	0.015	0.013	0.012
$s=1.0$	0.008	0.008	0.007	0.009	0.008
$s=1.5$	0.018	0.016	0.015	0.014	0.013
$s=2.0$	0.018	0.018	0.016	0.013	0.011
$s=2.5$	0.017	0.013	0.011	0.011	0.010
$s=5.0$	0.006	0.004	0.004	0.005	0.005
$s=7.5$	0.001	0.001	0.001	0.001	0.001
$s=10.0$	0.000	0.000	0.000	0.000	0.000
$s=15.0$	0.000	0.000	0.000	0.000	0.000
$s=20.0$	0.000	0.000	0.000	0.000	0.000

Table 8: Approximate Henze-Schilling Statistic values for two unit-normal distributions ($d=20$) with sample size of $n_1=n_2=500$, as a function of mean separation, for $k=\{1,2,3,4,5\}$, std. dev. for 10 realizations.

Conclusion

We evaluated three distribution-free, asymptotic measures of feature set divergence / separability. For low sample sizes ($n < 500$) and low dimensionality, the methods have similar behavior. For low sample sizes and higher dimensionality, the α Jensen Rényi Divergence appears to be less stable than the other two methods. The Henze-Penrose Divergence and the Henze-Schilling Statistic are found to have similar values. Currently our Henze-Penrose Divergence routine calls MST code written in *c*, but the Henze-Schilling Statistic routine is coded entirely in MATLAB. Once we rewrite the Henze-Schilling Statistic in *c*, we can evaluate relative runtimes.

Of the three methods tested against location alternatives for two unit-variance normal distributions, the Henze-Penrose Divergence and the Henze-Schilling Statistic give similar results that appear to have less bias and less variance at higher dimensionality and low sample number than does the α Jensen Rényi Divergence. Once the two distributions are completely separated, the Henze-Penrose Divergence and Henze-Schilling Statistic no longer increase with increasing distributional separation, however the α Jensen Rényi Divergence continues to increase with increasing distributional separation. This difference in behavior can dictate the choice of distribution-free divergence measurement, given the requirements of the task at hand.

It is our intention to evaluate the α Jensen Rényi Divergence, the Henze-Penrose Divergence and the Henze-Schilling Statistic over a broader set of conditions, including higher dimensionality, scale alternatives and different types of distributions for the two feature sets. Also it is our intention to evaluate the other distribution-free divergence measurements listed in the Introduction.

2.A.1.d. CADSP UCIR Evaluation Technical Support

There is currently a great deal of interest in UCIR sensors for Automatic Target Acquisition (ATA) on smart munitions, such as the NetFires NLOS PAM. The Georgia Tech CADSP imager has the potential for being incorporated into *on-Focal Plane Array* (FPA) pre-processing operations; these include: Non-Uniformity Compensation (NUC) and non-linear/non-local pixel equalization. Traditional equalization approaches (*e.g.*, histogram equalization) tend to perform very poorly and it is likely that a localized, non-linear equalization approach is needed. Given ISP Phase II funding constraints, we will

limit these pre-processing investigations to an evaluation of their implementation on the Georgia Tech CADSP imager. We have also been in preliminary discussions with Eglin, Air Force Base about using their optical flow test facilities.

2.A.2. ASU Technical Progress

2.A.2.a. Tracking and Sensor Scheduling with Motes Demonstration

We have continued to investigate sensor configuration for the ISP demonstration of tracking a target moving through a network of sensor motes. Specifically, we developed a myopic sensor scheduling algorithm that activates the lowest network energy cost combination of at most L sensors in the network to maintain a desired squared-error accuracy in the target's position estimate. The scheduling is performed using a linear approximation and the information formulation of the filter step of the Kalman filter. We pose the sensor scheduling problem as a discrete optimization problem with L binary-valued variables; each sensor has a corresponding binary variable that indicates whether the sensor will obtain a measurement in a given time epoch (a value of 1 implies that the sensor obtains a measurement). We have derived a 0–1 mixed integer programming (MIP) formulation for the case where the sensors provide only position information; we solve the MIP using a linear programming branch-and-bound algorithm. For the case when the sensors provide both position and velocity information, the scheduling problem is posed as a 0–1 convex program that we solve using the outer approximation algorithm. Our simulation results demonstrate that we can obtain optimal sensor scheduling for up to 50–70 sensors in the order of seconds.

We are continuing our characterization of the Mica-2/Z acoustic sensor response to a walking human target. We have conducted experiments to test the transmission and reception of acoustic data using the microphone on the sensor motes. We have implemented a sound recorder program that runs on the mote and continuously samples the data at 14 KHz but begins recording only when the samples have substantial energy. As many 8-bit samples as possible are collected; these are then transmitted to the base station.

We have resolved some packet corruption/loss issues and were able to achieve higher sampling rates than in our original experiments. For testing purposes, sinusoids with varying frequencies were generated and sensed by the microphone on the Mica-2 motes. The recorded data was transmitted in 36-bit packets to the base station and analyzed using Matlab. We observed fairly good reconstruction of the sinusoids for frequencies between 200 Hz and 5 KHz. We have also investigated the bandwidth of the acoustics signals generated by footsteps.

We collected data using a microphone connected to a PC; we have determined that human footsteps (tennis shoes on a hard floor) have a bandwidth of about 1-1.5 kHz. As sinusoids with such frequencies were successfully recorded and transmitted with the motes, we anticipate that footsteps can also be analyzed with this setup. Now that we are confident that the mote hardware is working correctly, the experiment of collecting acoustic data for a person walking in a circle around the sensor mote at various distances will be repeated to obtain accurate receiver operating curves and probability of detection versus threshold curves as well as an acoustic energy measurement model appropriate for our tracking and scheduling algorithms described above.

The acoustic model will be incorporated in a 1-bit tracker. Specifically, the energy detector will be processed at each sensor mote and the received energy will be compared to a set threshold in order to obtain binary decisions (detect/no-detect). The detector decisions encoded in a single bit will be communicated to a centralized processor that will use these measurements in a particle filter based tracker (implemented in Matlab) to estimate the target's position and velocity..

2.A.2.b. Multiple Target Tracking using the Configurable CADSP Imager

We continue to investigate the problem of tracking multiple targets in a surveillance region using image data from the CADSP imager. The objective is to be able to configure the imager to compute optical flow or image selective sub-areas of the field of view. We have successfully detected and estimated the motion of a single target (entering and leaving a scene) using synthetic data that we generated using the POV ray-tracing program. We have also simulated the use of a configurable imager that can provide only selective regions of interest in the imager as well as perform linear operations on these selected regions; the linear operations are Gaussian or Mexican hat filter operations that can be performed on video frames by the CADSP imager. Currently, we are experimenting with real data from a web-camera recording a person walking. After calibrating the camera and training the background and foreground distributions of a real scene, we will experiment with tracking the person entering and leaving a room; we will also extend these results to multiple persons.

2.A.3. UM Technical Progress

In the three months since the last quarterly report we have made progress on three fronts. First we have continued our development of classification constrained dimensionality reduction (CCDR) techniques for high dimensional classification problems. Second, we have developed new methods for localization in wireless sensor networks. Third, we have started to develop a new multivariate anomaly detection method that we call geometric entropy graphs (GEM) that is founded on our recently formulated theory of k-point minimal spanning tree (kMST) covering sets.

2.A.3.a. Progress on Classification Constrained Dimensionality Reduction

Progress has been made in the formulation of a low complexity out-of-sample extension of CCDR. In the revision of the paper [Raich and Hero 2006], accepted at the IEEE ICASSP conference after the last quarterly report, we established that a multi-class generalization of CCDR could outperform other algorithms in terms of average misclassification error on the benchmark LANDSAT dataset. The implementation of CCDR requires training (with cross-validation) a modified Laplacian Eigenmap (LE) dimensionality reduction algorithm to achieve an optimal tradeoff between over-fitting the LE manifold to the training data and misclassifying the training labels (using kNN or other simple classifier on the manifold).

Unlike PCA or local MDS algorithms, the LE classifier does not give an explicit mapping function from the measurement space to the classification space. Specifically, the LE mapping is implicit and its application to a new unlabeled sample, the so-called out-of-sample-extension, presents significant hurdles. In our ICASSP paper [Raich and Hero 2006] we used a heuristic method to perform this out-of-sample extension that requires application of the CCDR optimization to the union of the training samples and the new unlabeled sample. This method is impractical for online applications since it

requires growing memory and computation as new samples are taken into the system. Using a modified version of Nystrom's kernelization method as implemented for LE [Bengio et al. 2004] we can achieve an out-of-sample extension of CCDD that does not suffer from growing memory and computation. We are in the process of evaluating this out-of-sample extension for CCDD and will have more to report on it next quarter.

2.A.3.b. Progress on Localization in Sensor Networks

We have extended self-localization (calibration) methods and performance analyses to localizing transmitter nodes with either random, or completely unknown transmit power [Patwari and A. O. Hero 2006]. Transmit power calibration is the first step in solving the important problem of simultaneous tracking and calibration for wireless sensor networks. In realistic sensor networks, source energies and transmitter powers will vary by source and sensor, and over time. Power calibration eliminates any assumptions and makes localization and tracking robust to variation.

We have developed a new, low-complexity sensor localization algorithm, called the Laplacian Eigenmap Adaptive Neighbor (LEAN) algorithm [Patwari et al. 2006]. The LEAN algorithm and the Distributed Weighted MDS (dwMDS) method are both manifold learning methods that use pairwise distance measurements between nearby sensors. The LEAN algorithm has two main advantages. First, the LEAN algorithm solves directly for the global optimum. Secondly, it has low complexity compared to other globally-optimal algorithms. Both dwMDS and LEAN employ adaptive algorithms that improve robustness by de-emphasizing measurements that appear to be less reliable.

Further, we are currently testing new localization algorithms which use radio interferometric measurements (RIMs). Rather than estimate sensor coordinates directly from RIMs, we are testing an indirect method with two stages: (1) Estimate pairwise distances from RIMs; (2) Estimate sensor coordinates from pairwise distance estimates. The direct method has difficulty overcoming local optima if initialized too far from the global optimum. The promise of the indirect method is that neither stage of the process suffers significantly from local optima, and the combined computation is dramatically less intensive than direct solution. Distributed calculation also becomes possible. We have shown the feasibility of the indirect method and are evaluating its performance and robustness.

2.A.3.c. Geometric Entropy Graphs (GEM) for Anomaly Detection

Our motivation for this work was to counteract the severe sensitivity to spurious training samples of non-linear projection methods such as ISOMAP, LLE, HLE, and LE. For example, we have observed that the generalization error of CCDD suffers badly when a training sample has few near neighbors - such an anomalous point overly stretches the manifold to accommodate a smooth mapping. To control such errors these anomalous points must be identified so that they can subsequently be down-weighted or eliminated from the training sample. There has been substantial interest in identification of anomalous points in the machine learning community, notably the work on minimal volume sets by [Scott and Nowak 2005]. For Lebesgue continuous (no Dirac components) multivariate densities of the training data, the minimum volume set of specified coverage probability has several optimality properties in terms of anomaly detection. However, there are two problems with the approach of [Scott and Nowak 2005]: the minimum volume set loses its optimality properties when there are Dirac

components in the density; and the minimum volume set estimation procedure of [Scott and Nowak 2005] involves a relatively complicated set approximation based on multiresolution dyadic partitioning methods.

We have developed a theory of *geometric entropy manifolds* (GEM) which will likely overcome the two limitations of [Scott and Nowak 2005]. In GEM dimension reduction is combined with minimum entropy set estimation using ideas related to minimal spanning trees (MST), specifically approximations to the kMST [Hero and Michel 1999], i.e., the tree obtained by optimally pruning the MST (connecting all of the data points) to connect only a specified proportion $p=k/n$ of them, optimized to have minimal entropy. Points falling outside the kMST are then declared anomalies. This test is asymptotically optimal: It maximizes correct detection probability among all tests of specified false alarm level. Moreover, the minimum entropy set is naturally dimensionality reducing and reduces to the minimal volume set when the multivariate density of the training data is Lebesgue continuous. Figure 18 illustrates the GEM approach as compared with the approach of [Scott and Nowak 2005]. The points captured by the kMST, i.e. those connected by the red line, cover the most concentrated 90% of the points. The anomaly detector is implemented by inserting the new data point into the sample and re-computing the kMST. An anomaly is declared if the kMST does not span the new data point. The computation is entirely data driven, adapts to intrinsic data dimensionality, and does not involve explicit set estimation. Computational efficiency is still an issue, however, and methods for approximating the kMST for large data sets have been proposed using region of influence sets of the kMST. The mathematical and algorithmic foundations for such methods are currently being developed and a report is currently in preparation.

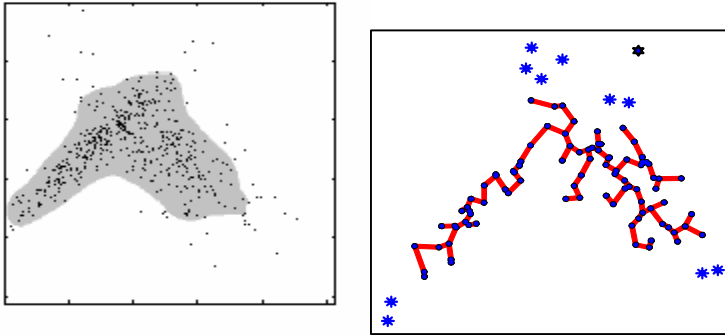


Figure 18: A simulation of a Gaussian mixture density in the plane. At left is the minimal volume set (gray) with coverage probability of 0.9 estimated by the multiresolution method of [Scott and Nowak 2005]. Any points falling outside of the minimum volume set would be classified as anomalous with level of significance 0.1. At right is the GEM method (illustrated for a smaller number of points from the same distributions at left) using the k-MST method. The minimum entropy set covering a specified proportion (0.9) of the probability mass of the distribution of the training samples is defined as the region of covered by the minimal spanning tree that covers at least 90% points. The asterisks denote training samples and the star denotes a potential anomaly. The coverage region can be defined either implicitly or explicitly using combinatorial optimization techniques.

2.A.4. FMAH Technical Progress

2.A.4.a. Nearest Neighbor Metrics

Over the last two months a new adaptive nearest neighbor algorithm has been developed. The goal for the work performed was to try to find a satisfying framework for the metrics that can be used to perform efficient approximation of the optimal Sensing Waveform- Design. Specific applications for these mathematical Waveform Processing methodologies are currently being identified.

Hybrid Nearest Neighbor Metrics

Previous attempts to address the problem of characterizing the Scattered Signal Space have been totally disconnected from the context within which the data is collected in practical applications. This means that the problem was often reduced to the one of return classification: e.g. by identifying a relationship between the horizontal and vertical range of movement of the objects – or other, more complex geometric relationships between time-frequency components.

However, all modern surveillance problems are entirely site/application specific so that a large amount of variability in the characteristics of data collected under different conditions has to be considered as a component of the problem. Thus, a substantial degree of freedom in feature selection is a crucial requirement for any reliable algorithm. At the same time *sidelobe reduction* problem is of fundamental importance in a variety of Radar applications.

This is why over the last eight weeks we started the development of a technique capable of reducing the Integrated Side-lobe Level (ISL) of a mixed signal r , consisting of a Target return, Multi-paths (attenuated and delayed) and Noise Clutter returns:

$$\hat{r}(t) = \sum_{\ell=1 \dots N} a_{\ell} p(t) + b_{\ell} s(t - \tau_{\ell}) + s(t) + n(t) \quad (1)$$

The idea is to find an optimal waveform design via a series of subsequent approximations. The assumption usually is that the Target components in (1) are stationary over a few duty cycles and that the noise component is Gaussian. Detection is performed by simply thresholding the output of the correlator.

Algorithm Overview

The main idea for new Nearest Neighbor Algorithm we developed is to combine features of the Laplacian Eigenmap [Belkin & Niyogi] algorithm with those of the Local Linear Embedding (LLE) introduced by L. Saul. For both processing methods, the fundamental assumption is that we can build a global, low dimensional parameterization of the data by determining the Nearest Neighbors of each data-point, provided that a sensible choice of the fixed number of the nearest neighbors for each point can be made. However, despite of their many successful applications, the algorithms mentioned above will fall short of being useful in a realistic application for the data we are interested in:

- Reliable estimates for the effect of noise or systematic distortions on the results of these algorithms are extremely complicated even in trivial cases (e.g. the very notion of a sufficient number of parameters cannot be determined a priori).

- Due to the specific character of the motion we are analyzing, it may be impossible to effectively separate the elements of different classes (e.g. distances between unrelated motions may be smaller than those between similar ones).

To overcome these obstacles we introduce a one-shot solution: we adaptively perturb the matrix of distances between the points by a small (i.e. bounded) amount in order to get a more appropriate local metric. This can be easily done as follows: the distances between points are iteratively modified in a Stochastic Gradient loop until the cost of the newly obtained metric configurations satisfies a predetermined, sub-optimal cost-criterion, say:

$$C(t) = F(s_1(t), \dots, s_N(t), \dots, m_1(t), \dots, m_M(t)) \quad (2)$$

The cost function in (2) simply measures the distance between the clusters of points corresponding to the different classes of received Signals (no-target, target, clutter etc.), after a set of waveforms s_k has been sent out and weighted with weight functions m_k .

2.A.4.b. Diffusion based high dimensional data processing

We begin with a brief discussion of a novel segmentation scheme based on our diffusion algorithm. This algorithm will be applied to uncooled infrared imagery that is of great interest to the NetFires NLOS program as well as several other multi-mode sensor programs that Raytheon is pursuing.

Overview

The goal of this portion of the project is to apply and adapt geometric diffusion methods of Coifman *et al.* to IR video data. As a simplification, we consider direct application of diffusion operators to the data itself.

Algorithm

Consider an image \mathbf{u} consisting of a rectangular array of pixels. We construct a diffusion filter \mathbf{K} , such that repeated application of \mathbf{K} to \mathbf{u} suppresses the background while enhancing or preserving regions of interest in the image. Furthermore, we *reinitialize* the filter after a fixed number of steps, using the output of the previous iteration to generate a new filter which is then applied iteratively to the output of the previous iteration.

Construction of \mathbf{K} proceeds as follows. For each pixel j in \mathbf{u} , a group of neighboring pixels of size $(2n+1) \times (2n+1)$, denoted $\mathbf{x}(j)$ and center on the pixel j , is selected. Next, a scalar non-negative kernel function G is selected to measure the degree of similarity or difference of two groups $\mathbf{x}(i)$ and $\mathbf{x}(j)$. The function G is therefore a bivariate function of vectors of length $(2n+1) \times (2n+1)$. If $\mathbf{x}(i) = \mathbf{x}(j)$, then $G(\mathbf{x}(i), \mathbf{x}(j)) = 0$, and if $\mathbf{x}(i)$ and $\mathbf{x}(j)$ are “dissimilar” then $G(\mathbf{x}(i), \mathbf{x}(j))$ should be large. Usually, G is designed to be a symmetric function of its arguments, so that the degree of similarity between $\mathbf{x}(i)$ and $\mathbf{x}(j)$ is the same as the similarity between $\mathbf{x}(j)$ and $\mathbf{x}(i)$.

From this kernel function, together with an integer N and a small parameter, ϵ , a matrix \mathbf{K}' is constructed so that

$$\mathbf{K}'_{ij} = \begin{cases} e^{-G(\mathbf{x}(i), \mathbf{x}(j))/\epsilon} & \text{if } |i-j| \leq R \\ 0 & \text{otherwise} \end{cases}.$$

Then, \mathbf{K} is constructed from \mathbf{K}' by normalizing the row sums to one; that is,

$$K_{ij} = K'_{ij} / \sum_l K'_{il}$$

where the sum is over the columns. The matrix \mathbf{K} may be viewed alternatively as a *diffusion operator*, since its action on \mathbf{u} is identical to a forward time step in a discrete diffusion problem, or as a *Markov matrix*, since the sums of its rows are unity. Each row has $(2R+1)^2$ non-zero entries, and the total number of rows is equal to the number of pixels in \mathbf{u} .

Initially, we utilize the square of the Euclidean distance to define

$$G(x(i), x(j)) = \|x(i) - x(j)\|^2,$$

a choice that, though apparently simple, is remarkably powerful. In principle, a wide variety of possible choices could be made for G .

Thus, the segmentation algorithm has the following steps:

Let $\mathbf{u}_1 = \mathbf{u}$. For $m=1, \dots, M$:

1. Define \mathbf{K}_m using input \mathbf{u}_m as above.
2. Compute $\mathbf{u}_{m+1} = (\mathbf{K}_m)^N \mathbf{u}_m$

The input to the algorithm is the original image \mathbf{u} . The parameters of the algorithm are

- G, ϵ .
- R, n . The radius of the filter window and the radius of the pixel sets used to form local neighborhoods in the image.
- M, N . The number of outer and inner iterations segmentation iteration.

Application to uncooled IR images

Below, we apply the algorithm to an uncooled IR image shown in Figure 19. The goal of the algorithm is to segment the image into “Target” and “Nontarget” regions. After the segmentation process, the segmented image could be used to pull out the target portion of the image for further analysis (such as classification/identification, *etc.*); for now, we focus only on the segmentation process.

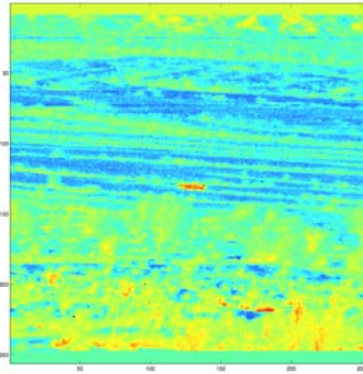


Figure 19: Single frame of uncooled IR data. The image contains both significant noise, as well as ground clutter.

Figure 20 shows the result of a few inner iterations of the algorithm. Though noise is suppressed, no segmentation has occurred. Increasing the number of inner iterations produces the results in Figure 21. Ground clutter is still somewhat evident. Finally, in Figure 22, much improved results are obtained by utilizing both inner and outer iterations. The algorithm shows remarkably robust performance, given the level of ground clutter and noise in the image. Edge detection methods applied to this same image, as seen in Figure 23 for example, produce results containing a large number of edges due to ground clutter.

Possible next steps

- Explore statistical or other functions in place of Euclidean distance for the choice of G . Different types of targets or images might benefit from other choices.
- Test performance of algorithm on images at a variety of ranges, with the goal of developing a 'training' set of parameters for a given target at close range, which can then be used for target segmentation/identification at longer ranges.
- Test algorithm on a wider variety of images, including uncooled IR with higher noise levels, lower contrast, and more ground clutter.
- Exploit interframe data in IR video.
- Develop algorithms for automated or semi-automated choice of parameters (currently parameters are selected through manual trial-and-error).
- Develop algorithms for post-processing results to eliminate spurious hot/cold spots, preserve target detail in original image, *etc.*

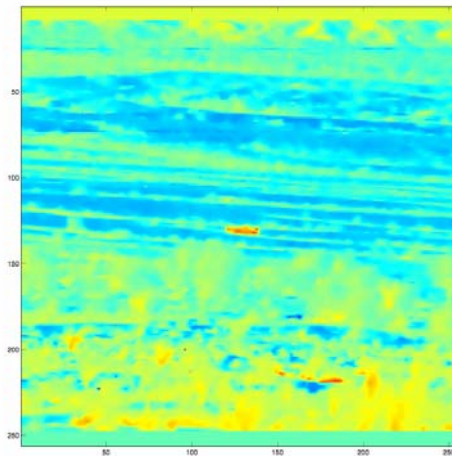


Figure 20: $R = n = 1$, $\epsilon = 0.0005$, $M = 1$, $N = 4$. In this image, noise has been reduced substantially, but no appreciable segmentation has occurred. In particular, the ground clutter is still quite prominent.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

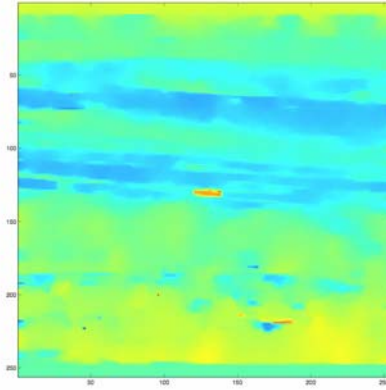


Figure 21: $R = n = 1$, $\epsilon=0.0005$, $M = 1$, $N = 100$. By increasing the number of inner iterations, clutter is reduced somewhat, but target detail has also deteriorated.

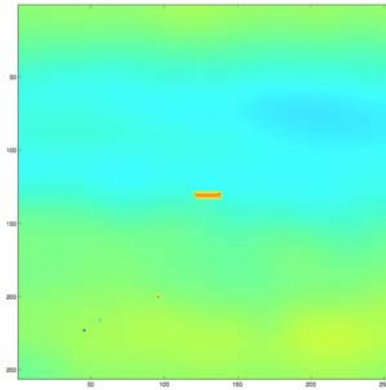


Figure 22: $R = n = 1$, $\epsilon=0.0005$, $M = 100$, $N = 4$. Instead of a large number of inner iterations, we use a large number of outer iterations. Target detail has been reduced, but the target's location is preserved. Ground clutter has been almost totally suppressed (with the exception of a few hot/cold spots).

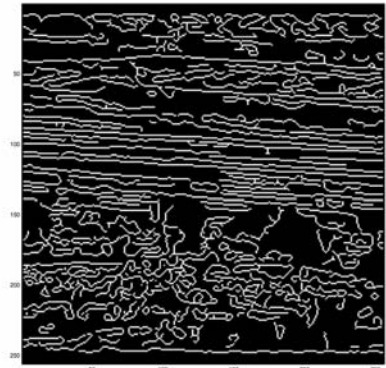


Figure 23: 'Canny' edge detection applied to original image from Figure 1. Though the target edge is indeed visible, ground clutter clearly presents a major obstacle to any segmentation algorithm based on edge detection.

The initial results appear to be very promising. The diffusion approach was able to homogenize the (difficult) clutter, while largely preserving the target. We are in the process of getting access to more realistic UCIR imagery, where there is significantly less target-background contrast as well as more varied and challenging clutter.

2.A.5. Georgia Tech Technical Progress

2.A.5.a Imager IC Development Status

This quarter design was completed for another generation of the analog processing separable transform imager. This version was design on a 0.35um process with a resolution of 256x256. The block transform size is 16x16 with a new ability of 8 pixel offsets allowing for overlaps of the regions of transformation. This extends the capability of the architecture beyond simple block transforms. In particular, this allows for more general 8x8 separable convolutions. Again, the imager is designed for separable transforms though more general transformations are possible with less then optimal performance, notably speed. This last version included two A matrices and two B matrices, though they may not be programmed while acquiring images.

Though not completed, progress was made toward the goal of run time programming, where one transform may be programmed while another is being used. In design and analysis, there where no particular obstacles found to achieving this goal, but some practical limitations, most essentially time, deterred placing that capability on this chip. At this point is believed the next version will have that ability. Most efforts in this design were concentrated on new circuits for reading wide ranges of currents and minimize sources of mismatch and error. Previously encountered stability issues were examined and changes were made accordingly. This includes architecture tweaks and novel circuit modifications. Included in this effort was the inclusion of automatic gain compensation amplifiers to work with wide ranges of currents over several orders of magnitude, even with wide ranging local image variances.

Currently development in board design and coding changes are in progress for testing the new chip. Initial tests will be done with image projections using optical test bench.

2.A.5.b Optical Flow Algorithm Status

The goal here has been to develop efficient algorithms for optical flow estimation that are closely tied to and accelerated by the CADSP imager. The algorithm on which we base this work is the LK-OFE (Lukas-Kanade Optical Flow Estimation) algorithm. The constraints placed by the CADSP imager architecture can be summarized as:

1. Only sequential signal flow is realizable.
2. One column- and one row-wise filtering operations are possible.
3. Matrix type operation for image is supported.

What we developed:

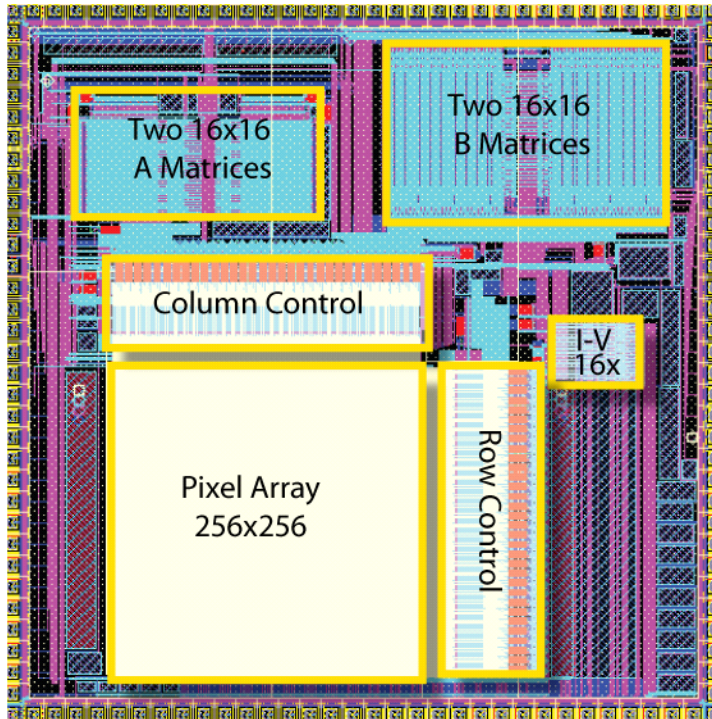
Checkerboard-type filtering for spatial derivatives
(Spatially sub-sampled filtering of OFE for derivative values (Ix, Iy))

Results

1. Only 2-stage filtering is necessary for spatial OFE derivative values (Ix, Iy).
2. Anti-aliasing operations are inherently performed for sub-sampling.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

3. Accuracy of spatially sub-sampled filtering of OFE is almost same with that of original LK-OFE for translating and diverging tree sequences.
4. One basis memory for each row- and column-wise filtering is enough.
5. Compared to the original sub-sampled LK-OFE, the number of operation and memory size for subsequent digital filtering operations (required as part of the LK-OFE) is somewhat reduced.
6. The density of our OFE is dropped between 3 and 3.5 times with respect to that of the original LK-OFE.
7. What we developed can be applied for any gradient-based OFE algorithm.



Future work

1. To increase the density of optical flow field, the interpolation of optical flow field would be proper.
2. Some simple post processing to improve the accuracy of optical flow field may be added.

2.A.6. UniMelb Technical Progress

2.A.5.a Raytheon Technical Support

Raytheon has been working closely with the University of Melbourne (UM). This collaboration is strengthened by the UM liaison, Craig Savage, working on a PhD with the UM personnel. During the past year, Raytheon and UM have worked together in three major areas:

1. Optimal scheduling for Gauss-Markov systems
2. Sensor scheduling for targeting/tracking applications
3. Waveform design for tracking

The first area is more theoretical, while the others have definite practical applications. We detail our work in the areas below.

Optimal Scheduling for Gauss-Markov Systems

In previous quarters, we have shown results for the optimal schedule for scalar systems undergoing a random walk. That is, for N systems, each system is evolving as

$$x_{k+1} = Fx_k + w_k$$

where $w \sim N(0, q)$ is a noise term. If each system's state is estimated by a Kalman filter, then we have shown optimal scheduling results. One interesting case is that if the only difference between states is the initial state covariance of each system, p_0^i , and each system may be measured at most once over a time horizon of length T , then the optimal schedule is to measure the T systems with the highest variances in increasing order. This is interesting as it is the opposite of the greedy case, which would measure systems with the highest variance first. Conversely, according to work by Howard, Suvorova and Moran [HSM], for a cumulative cost function, greedy appears to be optimal. Future work will include attempts to unify the results to yield a more general scheduling structure, with the aim to gain insight to the general problem of when a greedy schedule is optimal.

Furthermore, we have identified similar results for when the states of each systems are vectors when state estimation is performed with a fixed-gain filter. Our results indicate that, unlike the scalar case, the optimal schedule is a function of only the process noise covariance, Q , and the "tuning" of the fixed gain. These results are evident by the fixed gain equations, in that neither the current estimated state covariance nor the measurement error covariance appear in the formulae for estimated state covariance propagation. Our scheduling work has been submitted to the 9th International Conference on Information Fusion for publication.

Sensor Scheduling for Targeting and Tracking Applications

Emitter location using time difference of arrival (TDOA) techniques was the primary focus of our previous work. In TDOA applications, the relative geometry between the emitter and passive receivers plays an important role in targeting accuracy. Thus, our previous work on scheduling receivers amounted to selecting appropriate positioning for the receivers, either in selecting trajectories for the receivers (as in UAVs), or by choosing distributed receivers to collect measurements. This initial work assumed that the ability of a receiver to detect the emitter was independent of the geometry. While that assumption simplified the mathematics, it ignores the practical difficulty of substantial differences in beam patterns. The energy in the main beam is substantially greater than the power in the back lobe; hence, we would expect the probability of detection to vary accordingly. Raytheon is working with UM to define the problem appropriately, and to extend the solution to include multiple emitters. More information on our progress, including mathematical developments, may be found in the UM section of this report.

During the past quarter, we have developed a different targeting paradigm. We consider the case of a primary observer equipped with a laser rangefinder, GPS receiver and a compass. One potential targeting idea would be for the observer to use the rangefinder and compass to determine the target's polar coordinates relative to the

observer's position. However, such errors are dominated by the observer's ability to accurately determine azimuth angle with the compass, and these angle estimation errors impact the targeting cross-range solution proportionally with range.

If, however, there are passive secondary receivers capable of detecting reflected laser energy, these receivers may reduce uncertainty in the cross-range dimension, relative to the primary observer case. We have explored results for three cases:

1. Primary only, in which no secondary measurements are available
2. Fully observed, where the probability of detection of all receivers is 1
3. Partially observed, where the probability of detection of secondary receivers is 0.5

Example tracks for a target moving with nominally constant velocity of 10 m/s due east are indicated in **Figure 24**. The track in each case was computed using an unscented Kalman filter [Julier1]. Average state estimation performance is presented in **Figure 25**.

We are currently extending our work to include the scheduling of secondary receivers to improve track performance. This work leverages the TDOA scheduling algorithms researched in previous quarters, although the equations are different.

A paper regarding the general derivation of state estimation for the problem has been submitted to the 9th International Conference on Information Fusion. A second paper, including scheduling results, is being written for submission to the *IEEE Transactions on Aerospace and Electronic Systems*.

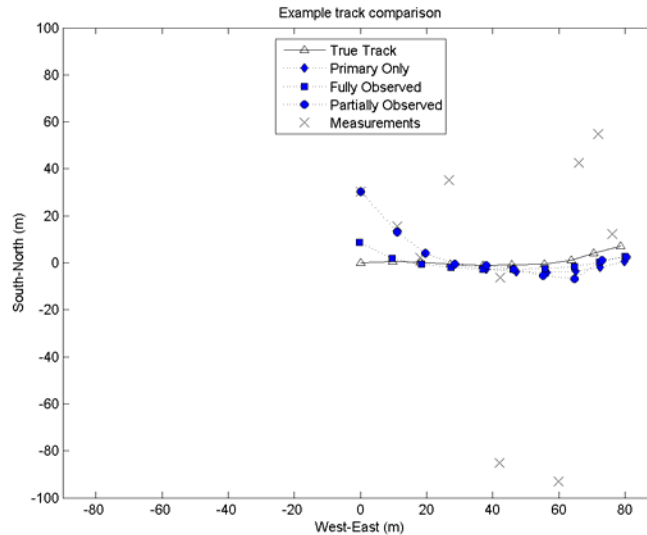


Figure 24: Comparison of tracks in primary only (diamond), fully observed (square), and partially observed (circle) cases. Primary only solutions (from range and azimuth information) are indicated by X. Track is estimated using an unscented Kalman filter.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

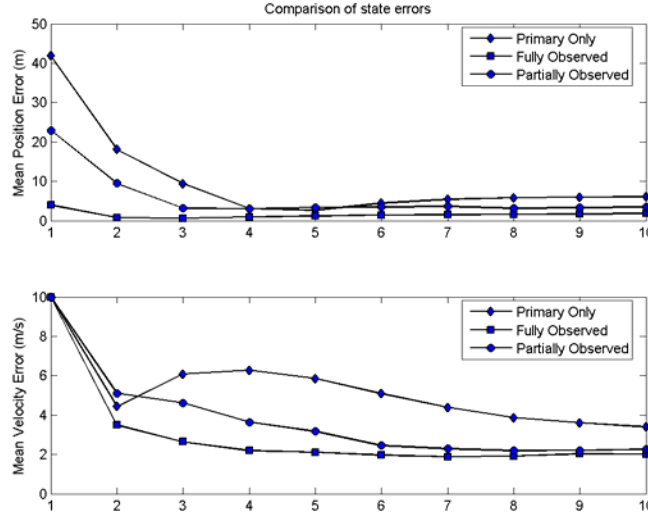


Figure 25: State estimation errors for 3 cases. For fully observed case, there is only a small amount of difference in position error of the system across multiple measurements.

Waveform Design for Tracking

In previous quarters, we examined waveform selection for tracking within the framework of an interacting multiple model (IMM) state estimation algorithm. The work was an extension of work undertaken during the waveform adaptive signal processing for missiles (WAS-M) program. The waveforms were selected by application of fractional Fourier transform (FrFT), which generates a rotation of a lower bound of the measurement error. The work and results were accepted for publication in the *IEEE Transactions on Aerospace and Electronic Systems*.

That work, while novel in scheduling for an IMM, assumed a pristine environment, namely one target characterized by a probability detection of one, with neither clutter nor false alarms. In practice, waveforms must attempt to mitigate problems with missing detections and/or clutter. While our work on scheduling for Gauss-Markov systems has had little to do with waveforms per se, the choice of waveform in tracking applications may be included in a Gauss-Markov system by allowing for a choice of the measurement noise covariance, R . Furthermore, recent work by Sinopoli *et al.* [Sino] and Gupta *et al.* [Gupta], have allowed for theoretical limits based on different measurement reception rates. While their results were derived in the context of imperfect measurement transmission/reception in sensor networks, their results are applicable in cases with a probability of detection, λ , less than one. In particular, their results indicate upper and lower bounds for the expected estimated state covariance as a function of λ . Such limits may be used to quantify a λ (probability of detection) vs. R (measurement covariance) trade-off.

Other future work includes designing waveforms to reduce clutter density. Reducing clutter is primarily a practical concern in that it is difficult to analytically determine tracker performance improvement as a function of clutter density reduction.

2.A.5.b Greedy versus Non-Greedy Scheduling

Minimum Entropy Scheduling for Hidden Markov Processes

In this project we have studied the fundamental model for scheduling of Hidden Markov processes. The complete description of this work is in [RSM], which has been submitted to the 2006 International Conference on Distributed Computing in Sensor Systems.

We generalize a Hidden Markov Process $\{S_n\}_{n=0}^{\infty}$ to be a process defined by $[S, P, T^k (k=1,2,\dots,K), Z]$, where S and Z are sets of possible states and measurement outcomes respectively, P is a transition probability matrix, and $T^k (k=1,2,\dots,K)$ are observation probability matrices. In contrast to the usual hidden Markov process [Eph] which has only one observation probability matrix, here the measurement Z_n at time n is related to the state S_n through the observation probability matrix T^{k_n} which varies with time index n . The aim is to find an optimal policy to chose the measurement sensor k_n (based on state estimate at time n) which provides the best observability of the states in the long run. This problem arises in applications for optimal usage of a set of sensors for observation of a Markov process where the system or the resource management allows only one sensor usage at a time. For example, in a radar system only one waveform out of a set can be used at each pulse transmission.

We denote Δ as the space of probability measures on the state space S and $P(\Delta)$ as the space of probability distributions on Δ . The *information state* π_n as a random variable on Δ , defined by $\pi_n(Z^{n-1}) = p(S_n | Z^{n-1})$, is a sufficient statistics for all the past measurements in relation to state estimation and is the basis for selecting sensors. Therefore a partitioning of Δ can be considered as a policy for sensor selection where we consider our selection based on state estimation through recursive update of $\pi \in \Delta$ from

$$\pi_{n+1} = f^k(z, \pi_n) = \frac{\pi_n D^k(z) P}{\pi_n D^k(z) \mathbf{1}} \quad (1)$$

In (1), $D^k(z)$ is a diagonal matrix with $d_{ii}(z) = T^k[i, z]$, $i=1,\dots,|S|$. The problem of finding the best policy has two major thrusts, defining a proper criterion by which a policy τ is assessed, and solving the optimization problem over different policies. Our achievements in these two parts are as follows.

Assessing policies

The entropy rate of a process is a measure of average information that each symbol carries. For a hidden Markov process the entropy rate is defined as $\lim_{n \rightarrow \infty} H(Z_n | Z^{n-1})$. Integral expressions for evaluating this entropy rate have been obtained in [Blkwell] and [Reza]. This entropy rate represents the ambiguity of the next measurement considering all past measurements which incorporates both uncertainty about state and the measurement noise in the next step. In order to have the best observability of the state we consider a measure which indicates uncertainty of state estimation from past measurements, and hence a suitable criterion for our scheduling

would be the limiting conditional entropy $\hat{H} = \lim_{n \rightarrow \infty} H(S_n | Z^{n-1})$. By a method similar to [Reza], we have obtained an integral expression for $\hat{H}(\tau)$

$$\hat{H}(\tau) = \int_{\Delta} h(\pi) \mu^{\tau}(\pi) d\pi \quad (2)$$

where $h(\cdot)$ is the entropy function and $\mu^{\tau} \in P(\Delta)$ is the stationary distribution of the Markov process $\{\pi_n\}_{n=0}^{\infty}$. The dynamics of this Markov process is governed by (1) where k is selected according to the policy τ . The stability of this Markov process [Tweem] is crucial in obtaining μ^{τ} and applicability of (2) which is still an open question. This stability ensures that there exists an operator Φ such that

$$\mu^{\tau} = \lim_{n \rightarrow \infty} (\Phi^{\tau})^n(\mu) \quad (3)$$

for any $\mu \in P(\Delta)$.

Optimal policy

We define a set of maps from Δ to $P(\Delta)$ as:

$$\phi^k(\pi) = \sum_z (\pi T^k)_z \delta(f^k(z, \pi))$$

for $k=1,2,\dots,K$, where $\delta(\pi)$ is a point in $P(\Delta)$ representing the probability distribution with all probability mass at π . The operator Φ^{τ} in (3) will be the integral of ϕ^k over Δ where in each region the index k is selected according to the policy τ . Based on (3) we have shown that the solution to the minimization problem

$$\tau^* = \arg \lim_{\tau} \hat{H}(\tau)$$

is

$$\tau^*(\pi) = \arg \lim_{j \in \{1,2,\dots,K\}} h(\phi^j(\pi))$$

where

$$\underline{h}(\phi^k(\pi)) = \int_{\Delta} h(u) d\phi^k(\pi)(u)$$

This solution shows that the optimal policy can be found only by evaluating a set of functions $\underline{h}(\phi^j(\pi))$, $1 \leq j \leq K$ over the domain Δ , and assigning to each region a sensor corresponding to a value k which gives the minimum of the functions $\underline{h}(\phi^j(\cdot))$ over index j .

Scheduling for IMM Filters

We assume that the dynamic models and the sensor measurement processes are linear and described by the following equations

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}(\Theta) \mathbf{x}_{k-1} + \mathbf{v}(\Theta) \\ \mathbf{z}_k &= \mathbf{H} \mathbf{x}_k + \mathbf{w}(\phi) \end{aligned}$$

where the dynamic model Θ is a discrete random variable, which at time k can take any value $1, \dots, M$ and $\phi = 1, \dots, N$ is the waveform used to obtain the measurements at time k .

A finite number of waveforms is considered. These waveforms form a *library*, as described in [SHME]. We write \mathbf{x}_k for the state of the track and \mathbf{z}_k for the measurement at time k . The matrices $\mathbf{F}(1), \dots, \mathbf{F}(M)$ are the state propagation matrices for the different maneuvers, \mathbf{H} is the measurement matrix. Process noise is denoted by $\mathbf{v}(1), \dots, \mathbf{v}(M)$ and measurement noise by $\mathbf{w}(1), \dots, \mathbf{w}(N)$. These are all zero mean, white, and uncorrelated Gaussian noise sequences with covariance matrices $\mathbf{Q}(1), \dots, \mathbf{Q}(M)$ and $\mathbf{R}(1), \dots, \mathbf{R}(N)$ respectively.

We assume that changes in target trajectory can be modeled as a Markov Chain with given transitional probability matrix \mathbf{P} , i.e.

$$\mathbf{P}_{i,j} = P(\Theta_k = j \mid \Theta_{k-1} = i) \quad i, j \in [1, M]$$

The trajectory of the target can be described at any time by one of the M dynamic models. The tracker switches modes between the dynamic models using the measurements, and thus facilitates tracking of maneuvering targets. Our problem is to choose the waveform which will minimize the entropy rate of dynamic model. We do this using a similar process to that used in Section 0 for scheduling of a hidden Markov process.

We denote the information state of the model at time k by $\mu_{k|k-1}(j) = P(\Theta_k = j)$ for $j = 1, \dots, M$. The likelihood function for a measurement given each model is

$$\begin{aligned} \Lambda_j^\phi(\mathbf{z}_k) &= P(\mathbf{Z} = \mathbf{z}_k \mid \Theta = j, \hat{\mathbf{x}}, \mathbf{Z}^{k-1}) \\ &= \frac{1}{\sqrt{2\pi} |\mathbf{S}(j, \phi)|} \exp\left(-\frac{1}{2}(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1})\mathbf{S}^{-1}(j, \phi)(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1})^T\right) \end{aligned}$$

where

$$\mathbf{S}(j, \phi) = \mathbf{H}\mathbf{P}_{k|k-1}(j)\mathbf{H}^T + \mathbf{R}(\phi)$$

is an innovation covariance matrix. We note, that \mathbf{S} explicitly depends on the waveform ϕ through the error covariance matrix $\mathbf{R}(\phi)$.

The posterior pdf of Θ is calculated using Bayes rule:

$$\mu_{k|k}(j) = \frac{\Lambda_j^\phi(\mathbf{z}_k)\mu_{k|k-1}(j)}{\sum_{i=1}^M \Lambda_i^\phi(\mathbf{z}_k)\mu_{k|k-1}(i)}$$

and $\mu_{k|k}(j)$ also depends explicitly on the waveform ϕ . Then the prior pdf of Θ (i.e. the information state at time $k+1$) is just

$$\mu_{k+1|k}(j) = \sum_{i=1}^M \mathbf{P}_{i,j} \mu_{k|k}(i)$$

Using the notation from Section 0, we can write

$$\mu_{k+1|k} = f^\phi(\mathbf{z}_k, \mu_{k|k-1})$$

where the measurements \mathbf{z}_k comes from an uncountable space. For a stationary policy $\tau = \{B_1, \dots, B_N\}$ we write

$$\mu_{k+1|k} = f^\tau(\mathbf{z}_k, \mu_{k|k-1}) = \sum_{\phi=1}^N f^\phi(\mathbf{z}_k, \mu_{k|k-1}) \chi(B_\phi)$$

where $\mu_{k+1|k} = [\mu_{k+1|k}(1), \dots, \mu_{k+1|k}(M)]$ and $\mu_{k|k-1} = [\mu_{k|k-1}(1), \dots, \mu_{k|k-1}(M)]$ are vectors. The decision function in this context is

$$h(\tau^*) = \min_{\tau} \int \sum_{i=1}^M \mu(i) \Lambda_i^\tau(\mathbf{z}) H(f^\tau(\mathbf{z}, \mu)) d\mathbf{z}$$

where $\sum_{i=1}^M \mu(i) \Lambda_i^\tau(\mathbf{z}) = P(Z = \mathbf{z})$ is the measurement likelihood, H is an entropy function and

τ is a policy. One can interpret this equation as the mutual information between the model and measurement

$$I(\Theta; Z) = \int P(Z = \mathbf{z}) \sum_{j=1}^M P(\Theta = j | \mathbf{z}) \log P(\Theta = j | \mathbf{z}) d\mathbf{z}$$

This explicitly suggests the our one-step ahead approach in the DASP paper [SHME] was, in this context, the optimum.

In DASP paper [SHME] we simply approximated $P(Z)$ by a single Gaussian but a more sophisticated method was developed for this work. It is described in [RSM].

2.A.5.c Tracking with Motes

Enhanced Tracking Performance with Signal Amplitude Information of Sensor Networks

In the last report, we presented a Virtual Measurement (VM) approach to the problem of tracking multi-target in a binary (motes) sensor network (see [WangMor]). In the VM approach, a set of activated sensor detections (with a binary measurement “1”) are mapped into a set of virtual measurements as if they were observed by a large sensor. The set of VMs can be straightforwardly fed into a classical multi-target tracker (MTT) for automatic multi-target tracking. A drawback of this method is that some VMs may have large variances when sensor nodes are sparsely distributed, which may cause large estimation errors. In this work, we have shown that the uncertainty of VMs that associated with 2 or more sensor detections can be reduced by using signal amplitude information. To enable this work, we assume that signal amplitude information from activated sensors is available to the base station (by transmitting an extra scalar number from each sensor to base station).

For a single target¹, received signal amplitude at the sensor is assumed to be exponentially decaying function of the relative distance between sensor and the target. In the absence of noise and calibration error, we can exactly localize the target location using the signal amplitudes detected by 3 or more sensors via a root finding algorithm. Even with the signal amplitudes detected by only 2 sensors, we can still calculate the location of the target at the direction in line between these two sensors. In the real world, these calculated locations come up as random variables and their uncertainties depend on the noise level (in terms of SNR) and sensor calibrations, etc. In the VM framework,

¹ In the VM approach, an IASG (Independently Activated Sensor Group) is deemed to have been triggered by a single target.

these calculated locations are treated as alternative VMs. The MTT will use the VM with smaller uncertainty.

Simulation results indicate that an improved tracking performance over the standard VM approach can be achieved by taking signal amplitude information into account when a detected signal of relatively high SNR is available. The proposed method is particularly useful when sensors of the network have large sensing ranges and sparse placement.

The uncertainty reduction of a VM that is due to a single sensor detection via signal amplitude information, and a better variance approximation for the variance of the VM calculated using signal amplitude information is currently being investigated. Full details of this new algorithm can be found in [WangMus].

Extensions to UKF-based Mote Tracker

An algorithm for tracking a target moving through a field of motes using a Gaussian mixture based approach was described in previous reports. This algorithm is based on using the unscented transformation to approximate the Kalman filter recursion. The algorithm was restricted to tracking a known number of targets with predictable dynamic behaviour, i.e., no maneuvers. This algorithm has been extended to include estimation of the target strength, target maneuvers and initiation of new tracks. Target strength is estimated by augmenting the target state to include a target strength parameter which is then recursively estimated along with the target kinematic parameters. Target maneuvers are handled by hypothesizing multiple motion models. The explosion in the number of possible motion model sequences is handled by pruning unlikely motion models at each time step. Interestingly, the popular interacting multiple model approximation seems to perform quite poorly with this model. Since mote detections signal the presence of targets, these are used to initiate new tracks. The validity of a track is measured by the probability of target existence which can be computed recursively as measurements are acquired. Tracks are confirmed as target tracks or deleted based on the existence probability.

Comparison of Mote Tracking Algorithms

Comparisons between the unscented Kalman filter (UKF) and another algorithm we have developed, the virtual measurement filter (VMF), have been performed. The algorithms assume different measurement models. The measurement model used by the UKF assumes that detection probability of a particular sensor decreases with target range while the model used by the VMF assumes that detection probability is fixed within the sensing range. Comparisons are performed with both measurement models to ensure equity. When a filter is used with a mis-matched measurement model it is necessary to select appropriate parameters for the measurement model assumed by the filter. Let $p(r; \theta)$ denote the detection probability for a target at a distance r from the sensor. The detection probability is parameterized by the vector θ . Assume that data is generated according to $q(r)$. We then select the parameter θ as

$$\theta^* = \arg \max_{\theta} \int_0^R [p(r; \theta) - q(r)]^2 dr \text{ subject to } p(0; \theta) = q(0)$$

To give an example, suppose that data is generated according to the model assumed by the VMF. Here

$$q(r) = \begin{cases} P_D, & r \leq S, \\ 0, & \text{otherwise.} \end{cases}$$

where P_D is the detection probability and S is the sensing range. The UKF assumes a model of the form

$$p(r; \tau, A) = 1 - \Phi\left(\Gamma - A \exp\left\{-\frac{r^2}{2\tau^2}\right\}\right)$$

where Φ is the distribution function of a standard Gaussian random variable and Γ is a threshold. Here $\theta = (\tau, A)$. According to the given criterion we set

$$A^* = \Gamma - \Phi^{-1}(1 - P_D)$$

and select τ to minimize the

$$\tau^* = \arg \max_{\tau} \int_0^R [p(r; \tau, A^*) - q(r)]^2 dr$$

An example of this fitting procedure is given in **Figure 26** for a detection probability of $P_D=0.7$ and a sensing range of $S=10$.

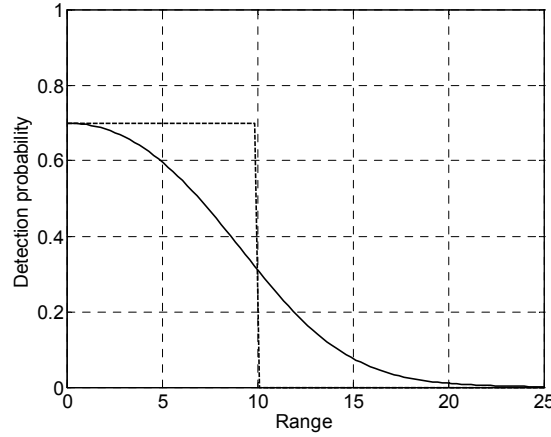


Figure 26: Detection probability plotted against target range for the UKF measurement model p (solid) and the VMF measurement model q (dashed).

The simulation set-up follows. The observation region is 120m x 120m containing 100 motes arranged on a rectangular grid. A range of sensor parameters is considered. In particular, data is generated according to the VMF model with all combinations of the detection probabilities $P_D = 0.5, 0.7, 0.9$ and sensing ranges $S = 5\text{m}, 10\text{m}$ and 15m . For each combination of P_D and S , data is generated using the UKF measurement model with parameters τ and A calculated as described above. Thus there are 18 data sets in total, 9 for each measurement model. The target moves with a velocity subject to low intensity Brownian motion. The algorithms begin with no knowledge of the location or number of targets. Tentative tracks are deemed true or false depending on their proximity to ground truth. A tentative track is confirmed once the existence probability, discussed above, exceeds a threshold. The aim is to confirm the true target track as quickly as possible and maintain while minimizing the number of confirmed false tracks.

The algorithm performances over 100 realizations are summarized in **Table 9** to **Table 12**. **Table 9** and **Table 10** show the results for measurements generated according to the model $p(r)$ defined above. **Table 9** shows the mean number of true track

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

confirmations averaged over time and measurement realizations. This number should be as close to one as possible. **Table 10** shows the mean number of false track confirmations averaged over time and the measurement realizations. This number should be close to zero. **Table 11** and **Table 12** display the same information for measurements generated according to $q(r)$. Each cell of the tables shows results for both the UKF and the VMF with the UKF results on the left. In all of the scenarios considered here the UKF outperforms the VMF, in terms of both the number of true tracks confirmed and the number of false tracks confirmed. With the exception of the case where $P_D=0.5$, $S=5$, the mean number of confirmed true tracks is close to one indicating that the true target track is quickly found and maintained. Also, no false tracks are confirmed. Only a slight drop in performance can be observed when measurements are generated with a mis-matched measurement model (recall that the UKF assumes a measurement model of the form described by p). Interestingly, in some cases the VMF performs better with a mis-matched measurement model, e.g., when the detection probability is 0.5 and the sensing range is 5.

Table 9: Mean number of confirmed true tracks for the UKF (left) and VMF (right) trackers. Data is generated using p .

P_D	Sensing Range					
	5		10		15	
0.5	0.65	0.23	0.93	0.55	0.97	0.81
0.7	0.86	0.64	0.96	0.83	0.97	0.83
0.9	0.91	0.79	0.98	0.83	0.98	0.88

Table 10: Mean number of confirmed false tracks for the UKF (left) and VMF (right). Data is generated using p .

P_D	Sensing Range					
	5		10		15	
0.5	0	0	0	0.02	0	0.01
0.7	0	0.01	0	0.01	0	0.01
0.9	0	0.01	0	0	0	0

Table 11: Mean number of confirmed true tracks for the UKF (left) and VMF (right). Data is generated using q .

P_D	Sensing Range					
	5		10		15	
0.5	0.64	0.13	0.93	0.59	0.97	0.78
0.7	0.84	0.57	0.97	0.83	0.98	0.81
0.9	0.88	0.77	0.98	0.88	0.96	0.88

Table 12: Mean number of confirmed false tracks for the UKF (left) and VMF (right). Data is generated using q .

P_D	Sensing Range					
	5		10		15	
0.5	0	0	0	0.10	0	0.02
0.7	0	0	0	0.38	0	0.03
0.9	0	0	0	0.83	0	0.04

The comparison performed here is limited in the sense that the target follows a straight line trajectory and only one target is considered. The comparison will be extended to consider initiation and tracking of multiple maneuvering targets.

Scheduled Tracking with Motes

In previous reports a single-target tracking algorithm using binary proximity sensors was described, where some sensors may be inoperative. Conventional methods assume all motes in the surveillance region are functioning. These trackers therefore assume that the lack of a target detection means that the target is not near the sensor, leading to biases in the track estimates. To track robustly with such sensors it is necessary for the central processor to know which motes in the surveillance region are functioning. The tracker described previously, called the **Health Tracker**, simultaneously estimates the target track and also the probability each mote in the surveillance region is operational. This algorithm is robust to sensor failures and gives improved track accuracy over conventional trackers. A complete description of the algorithm is given in [LaSMORESAV] and will be presented at the IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP) in Toulouse, France in June 2006.

This tracking approach has been extended to consider the case when the central processor can either listen for detections from the sensors or else query a particular mote to determine if it is operative. Querying a mote for its status results in a loss of information from the other motes, but certain information about the selected mote. This raises the question of when to schedule a mote query and which mote to select. Recent work investigated two such scheduling methods. One method was based on a *ad hoc* cost function while the other considered the expected cost trade-off. Simulation results show that both methods provide a moderate improvement in track accuracy compared to the basic **Health Tracker** approach and a significant decrease in computational complexity. A full description of this work is given in [LaS], which has been submitted for presentation at the 9th International Conference on Information Fusion in Florence, Italy in July 2006.

2.A.5.d Particle Filter Tracker for EKV

A model of the EKV problem has been formulated and a solution based on particle filtering has been developed. The EKV problem involves guiding one exo-atmospheric missile into another exo-atmospheric missile. The problem is complicated by the fact that the target missile deploys decoys. The goal is to ignore the decoys and focus on the target. An onboard algorithm is capable of classifying, with some degree of error, the various objects as targets or decoys.

The problem is mathematically described using a stochastic dynamic system. The state contains the positions of all objects in spherical coordinates with respect to some inertial frame of reference. The state also contains a discrete-valued variable which assumes the value of the object deemed to be the target. At each time step the missile provides object measurements in its own frame-of-reference. Each measurement is provided with a [0,1]-valued variable indicating the probability that it was generated by the target. The missile has a certain field-of-view (FOV) which must be manipulated, through the application of accelerations in directions orthogonal to the line between the missile and the target, to ensure that the target remains in the FOV. Clearly, as time goes

by the number of objects in the FOV will decrease and it is of particular interest that the target be in the FOV as the missile nears the target. The measurement model allows for uncertainty regarding the origin of each measurement, i.e., we don't know which measurement is due to which object.

It is proposed to select the acceleration inputs to maximize the probability that the target is in the FOV. Calculation of this probability requires knowledge of the joint posterior density of the state and the target indicator variable. Since this cannot be computed exactly it is proposed to use a particle filter (PF) to approximate it. The main problem is that the PF cannot be used to perform inference on static variables such as the target indicator. We therefore suggest using a PF to approximate only the density of the state conditional on the measurements. This results in a number of state trajectories. The posterior distribution of the target indicator variable can then be computed conditional on each state trajectory and the measurement history. The statistics required for the selection of the acceleration input can then be approximated. This filter will be implemented and tested using simulated data in the near future.

2.A.5.e Other Tracking Algorithms

Measurement Gaussian Sum Mixture Target Tracking

Details of this work are given in [MusEvans], which has been submitted to the 9th International Conference on Information Fusion (pending approval from the DARPA ISP project office). This method applies to target tracking when the measurement probability density function may be described or approximated by a Gaussian Mixture. Both single scan tracking (IPDA) and multi scan tracking (ITS) using this model are described. Extension to multi-target tracking via the Linear Multitarget procedure is straightforward, however it is not part of the submission.

Certain measurement non-linearities can be efficiently modeled as a Gaussian sum, and in these situations this tracker may present a viable alternative to using extended Kalman filter, particle filter or unscented Kalman filter. The paper itself describes 3 such situations:

- a) Tracking using acoustic signals with just two sensors; using signal amplitudes;
- b) Tracking using acoustic signal with just two sensors, using both signal amplitude and time difference of arrival; and
- c) Tracking using long range radars.

Situations a) and b) are relevant to the problem of tracking with motes. Time difference of arrival results in a similar measurement curve to the measurement function in scenario a). Passive radar sensors also have a measurement function that is significantly nonlinear. Future work will include investigating if these nonlinearities can be adequately modeled by a Gaussian sum, making this tracker relevant to such applications also. Application c) corresponds to long range radar, which may be used to track TBM at significant distances.

Tracking a Large Number of Targets in Clutter with Particle Filter

This work is described in [MusMore], which has been submitted to the European Signal Processing 2006 conference, (also pending approval from the DARPA ISP project office). It describes tracking a large number of targets simultaneously (50 targets) using

particle filters. The particle filter is obtained by applying Linear Multitarget method to the particle filter implementation of IPDA, previously published at the IEEE Conference on Decision and Control (CDC 2005) in Seville, Spain in December 2005.

The resulting algorithm is computationally efficient, at least compared to other particle filter implementations, and can be potentially applied for real-time target tracking. The majority of papers published on target tracking with particle filters demonstrate their performance on a small number of targets, typically as low as two or three. An additional benefit of this tracking algorithm is its ability to incorporate nonlinear target motion and measurement models due to the use of a particle filter. A tracking algorithm with these features is a necessary tool on which to base scheduling algorithms for problems such as sensor scheduling for swarms.

2.A.5.f Distance Preserving Projections

In [LaSMor] the problem of computing waveform measures of effectiveness was discussed. When performing waveform scheduling it is necessary to have some method for evaluating the effectiveness of any given waveform in the current environment. Four properties that such measures should satisfy were proposed. These were:

1. The measure should be rigorous, so it can be scientifically validated.
2. The measure should be computable.
3. The measure should be based on the total ambiguity of the waveform.
4. The measure should take into account the clutter distribution.

In [LaSMor] a minimum variance measure, the *MVMoE*, was proposed which satisfies properties 1, 3 and 4. A method for computing it was also given but it was computationally complex. To avoid the computational complexities it is proposed to investigate the use of distance preserving projections to determine if a more efficient means of computing the *MVMoE* can be found.

Distance preserving projections, also known as random projections, are based on a result by Johnson and Lindenstrauss [JohnLin] which shows that it is possible to project points in a high dimensional space into a lower dimensional space in such a way that all pairwise distances between the points are maintained to within an arbitrarily small factor. The projection matrices that achieve this are generated by selecting the matrix elements from simple random distributions. Examples of such projection matrices are given in works such as [Ach].

These projections may prove an effective means of calculating the *MVMoE*. The raw measurement used in the *MVMoE* is a vector of size MN where M is the number of range cells and N is the number of Doppler cells. For practical problems this vector is in a very high dimensional space. The evaluation method described in [LaSMor] required calculating the singular value decomposition of each waveform's ambiguity function to map this raw measurement vector to a lower dimensional space in order to compute the measure of effectiveness for each waveform. The use of random projections to perform this mapping should provide significant computational savings while preserving the relative value of the measure for each waveform.

2.A.5.g Sensor Scheduling for AT3 Applications

In this project we are examining sensor scheduling methods for geolocation of ground-based emitters using passive, airborne sensors. The initial scenario considered has

a number of simplifying assumptions that will be relaxed in later work. A description of the initial problem scenario is given below.

The targets are stationary, ground-based emitters operating at fixed, distinct frequencies. Each emitter is a phased-array radar whose boresight is fixed for the duration of the engagement. The passive sensors are mounted on airborne platforms, with known locations. Target geolocation is performed using time difference of arrival (TDOA) techniques. We assume that there is no pulse ambiguity and all the sensors are time synchronized. In this initial scenario, we assume that the receivers are evenly spaced around the perimeter of the region of interest and do not move significantly. That is, they may circle around some predetermined point but do not move noticeably in relation to the other sensors or the target. This assumption will be relaxed in later scenarios.

The sensors are assumed to be sufficiently sensitive that they are able to detect pulses from the emitters when they are in the sidelobes and backlobe of an emitter. A simple model for the probability of pulse detection, P_D , as a function of angle from boresight with three levels will be used – high P_D when near boresight; moderate P_D when further off boresight; and low P_D when the sensor is behind the emitter. The scheduling algorithms will be designed so they do not depend on the specific properties of the pulse detection model.

The sensors are assumed to be able to detect pulses over a wide frequency band – wide enough that all possible emitters can be detected. The sensors are assumed to be able to operate in one of two modes – *surveillance* mode and *tracking* mode. In *surveillance* mode the sensor is able to detect pulses on a relatively wide frequency band but the TOA measurement is relatively inaccurate. In *tracking* mode the detection range is narrower and the TOA accuracy higher.

It is assumed that some *a priori* knowledge of the number of emitters, their operating frequency and their location is available. The scheduling problem is then to determine which sensors should listen on what frequency band and in which mode in order to geolocate all the emitters as accurately as possible. Note, the emitters are not necessarily transmitting at all times during the engagement. Initially, myopic (i.e. greedy) scheduling algorithms will be developed, but we will also consider if more sophisticated schedulers can be designed for this problem given its particular structure.

2.A.5.h Scheduling for Passive Radar Sensors

This evaluation involves tracking stationary or slow moving (ground based) radar sources using UAVs, which measure Time Difference of Arrival (TDOA) of radar signals. Based on the TDOA values, the location of the radar can be estimated and tracked. For a 2 D situation, such as when tracking ground emitters, at least 3 UAVs are necessary. It is assumed that the UAVs may receive radar signal from the main beam, which is 1 degree wide, as well as from sidelobes. The radar antenna pattern will be approximated by assuming that, when not in the radar main beam, the radar signal is received 20% of the time. Each radar is assumed to have different carrier frequency, thus the signal origin is unambiguous. It is also assumed that the distance between the UAVs is smaller than the radar maximum range, thus there will be no pulse to pulse ambiguity.

The UAVs communicate using low frequency link, and the radar emits pulses with frequency of 5 kHz (medium range radar). Information about all pulses received can

not be transmitted. Thus the scheduling issue is the choice of the subset of received pulses, whose transmit time of arrival will be transmitted. The corresponding issue is the estimate of the radar position, given the received data from all UAVs.

2. B. Publications

There were no refereed publications that occurred during the current PoP.

1. Craig O. Savage and Bill Moran, "Waveform Selection For Maneuvering Targets Within An IMM Framework," IEEE Trans AES, accepted for publication.
2. A. Chhetri, D. Morrell and A. Papandreou-Suppappola, "Non-myopic sensor scheduling and its efficient implementation for target tracking applications," EURASIP Journal on Applied Signal Processing, to appear 2006.
3. A. Chhetri, D. Morrell and A. Papandreou-Suppappola, "On the use of binary programming for sensor scheduling," IEEE Transactions on Signal Processing, submitted February 2006.

2. C. Conference Proceedings

1. I. Kyriakides, D. Morrell and A. Papandreou-Suppappola, "Sequential Monte Carlo methods for tracking multiple targets with stochastic kinematic constraints," invited to the First IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, Puerto Vallarta, Mexico, December 2005.
2. W. Moran, C. O. Savage, S. Suvorova, H. A. Schmitt, D. E. Waagen and R. Cramer, "Dynamic Positioning and Scheduling of UAVs for Passive Geolocation," Session on cooperative dynamic systems, 2006 IEEE International conference on Networking, Sensing and Control, Ft. Lauderdale, FL, April 2006, accepted for publication.
3. S. Sira, A. Papandreou-Suppappola and D. Morrell, "Characterization of waveform performance in dynamically configured sensor systems," invited to the International Waveform Diversity and Design Conference, Kauai, Hawaii, January 2006.
4. R. Raich, J. A. Costa, and A. O. Hero III, "On Dimensionality Reduction for Classification and Its Application," 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, submitted.
5. B. F. La Scala, M. Morelande, C. O. Savage, "Robust Target Tracking with Unreliable Binary Proximity Sensors," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), submitted.
6. A. Chhetri, D. Morrell and A. Papandreou-Suppappola, "On the use of linear integer programming for sensor scheduling in sensor networks," submitted to the 5th International Conference on Information Processing in Sensor Networks, Nashville, TN, April 2006.
7. C. Vossberg, A. Swain, S. Bellofiore, B. Manjunath, D. Chakraborty, A. Chhetri, D. Morrell and A. Papandreou-Suppappola, "Sensor network prototype for tracking and scheduling with minimum resources," submitted to the IEEE Workshop on Sensor Array and Multi-channel Processing, July 2006.

8. A. Chhetri, D. Morrell and A. Papandreou-Suppappola, "Sensor scheduling using 0-1 mixed integer programming framework," submitted to the IEEE Workshop on Sensor Array and Multi-channel Processing, July 2006.

2. D. Consultative and Advisor Functions

There were two consultative or advisory functions that occurred during the current PoP. The first relates to a Raytheon Shooter Localization demonstration using the MICA-2/Z sensor nodes. This work is being funded under the DARPA IXO NEST Phase II program. The Phase I shooter localization algorithms were developed by VU. Preliminary results indicated that the shooter localization algorithm has significant potential. The program was subsequently classified and was ultimately transitioned to Raytheon for demonstration and refinement under Phase II. The DARPA IXO Program Manager has kindly given permission for several of these algorithms to be used in our ISP Phase II program. The Raytheon NEST program has identified a critical need for the development of an accurate sensor localization algorithm that is scalable to hundreds or thousands of nodes. Indeed, the DARPA NEST program hopes to demonstrate a 10,000 node network. We have identified and are evaluating several promising mathematical approaches to sensor localization developed by Al Hero (UM); these will be made available to the Raytheon NEST program if they are successful. Thom Steven and Sal Bellofiore support the DARPA ISP II and DARPA NEST programs, and, more generally, the two programs are developing a strong collaboration.

The second function relates to optical flow test facility at Eglin, Air Force Base. Raytheon and Georgia Tech have had preliminary discussion with Dr. T.J. Klausutis of Eglin AFB about the possibility of using their facility to evaluate the Georgia Tech CADSP imager being investigated on our ISP Phase II program. While these discussions are preliminary, Dr. Klausutis was interested in learning more about the capabilities and maturity of the CADSP Imager and plans to visit Georgia Tech.

2. E. New Discoveries, Inventions or Patent Disclosures

There were no patent disclosures filed during the current PoP.

2. F. Honors/Awards

There were no honors or awards received during the current PoP.

2. G. Transitions.

There were no technology transitions achieved during the current PoP.

2. H. References

[Aslan & Zech 2005] B. Aslan and G. Zech (2005). *Nuclear Instruments and Methods in Physics Research A* **537** 626.

[Baringhaus & Franz 2004] L. Baringhaus and C. Franz (2004). *Journal of Multivariate Analysis* **88** 190.

[Dwyer & Squire 1993] R. A. Dwyer and M.B. Squire (1993). *North Carolina State University Tech. Report TR-93-21*.

[Ferber 2000] D. Ferger (2000). *Journal of Multivariate Analysis* **74** 1.

[Friedman & Rafsky 1979] J. H. Friedman and L.C. Rafsky (1979). *Annals of Statistics* **7** 697.

[Henze 1988] N. Henze (1988). *Annals of Statistics* **16** 772.

- [Henze & Penrose 1999] N. Henze and M.D. Penrose (1999). *Annals of Statistics* **27** 290.
- [Hero et al. 2002] A. O. Hero, B. Ma, O. Michel and J. Gorman (2002). *IEEE Signal Processing Magazine* **19** 85.
- [Neemuchwala & Hero 2005] H. Neemuchwala and A. Hero (2005). *Proceedings SPIE* **5674** 99.
- [Rousson 2001] V. Rousson (2002). *Journal of Multivariate Analysis* **80** 43.
- [Schilling 1986] M. F. Schilling (1986). *Journal of the American Statistical Association* **81** 799.
- [Székely & Rizzo 2004] G. J. Székely and M.L. Rizzo (2004). *InterStat*.
- [Wei & Lachin 1984] L. J. Wei and J.M. Lachin (1984). *Journal of the American Statistical Association* **79** 653.
- [Raich and Hero 2006] R. Raich and A.O. Hero, "On dimensionality reduction for classification and its application," 2006 IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, Toulouse France, 2006.
- [Bengio et al. 2004] Y. Bengio, O. Delalleau, N. Le Roux, J.-F. Paiement, P. Vincent and M. Ouimet, "Learning eigenfunctions links spectral embedding and kernel PCA," *Neural Computation*, 16:10, pp: 2197--2219, 2004.
- [Patwari and A. O. Hero 2006] N. Patwari and A. O. Hero, "Source and sensor localization without known transmit power," (in preparation).
- [Patwari et al. 2006] N. Patwari, A.O. Hero, and J. Costa, "Learning sensor location from signal strength and connectivity," Chapter in edited volume on localization and time synchronization in sensor networks by Radhakrishna Sampigethaya, Springer-Verlag, 2006.
- [Scott and Nowak 2005] C. Scott and R. Nowak, "Learning minimum volume sets," NIPS 2005.
- [Hero and Michel 1999] A. O. Hero and O. Michel, "Asymptotic theory of greedy approximations to minimal K-point random graphs," *IEEE Trans. on Info Theory* , Vol. IT-45, pp. 1921-1939, 1999
- [Maroti, et al. 2005] Maroti, et al., "Radio Interferometric Positioning," Department of Mathematics, Vanderbilt University, *Technical Report* ISIS-05-602, 2005.
- [Costa et al 2006] J. Costa, N. Patwari, and A. Hero, "Distributed Multi-Dimensional with Adaptive Weighting for Node Localization in Sensor Networks," *IEEE/ACM Transactions on Sensor Networks*, to appear, 2006.
- [Meertens 2005] L. Meertens, "The Dimension of the Vector Space Spanned by Sets of Radio-Interferometric Measurements," Kestrel Institute, *Technical Report* KES.U.05.02, July, 2005.
- [Patwari 2006] N. Patwari, "Pair-wise, Distance-based Radio Interferometric Localization," *draft*, February, 2006.
- [HSM] S. Howard, S. Suvorova and B. Moran. "Optimal policy of scheduling of Gauss-Markov systems" *7th International Conference on Information Fusion*. Stockholm, Sweden, Vol. II, pp. 888-892, June, 2004.

- [Julier1] S. J. Julier and J. K. Uhlmann. "Unscented filtering and nonlinear estimation" *Proceedings of the IEEE*, Vol. 29, No. 3, pp. 401-422, 2004.
- [Sino] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan and S. Sastry, "Kalman filtering with intermittent observations", *IEEE Transactions On Automatic Control*, Vol. 49, No. 9, pp. 1453-1464, 2004.
- [Gupta] V. Gupta, T. Chung, B. Hassibi and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage", *Automatica*, Vol 42, No 2, pp. 251-260, 2006.
- [RSM] M. Rezaeian, S. Suvorova and B. Moran, "Sensor Scheduling for Observation of a Markov Process", Submitted to *International Conference on Distributed Computing in Sensor Systems*, San Francisco CA, USA, June 2006.
- [Blkwell] D. Blackwell, "The entropy of functions of finite-state Markov chains", *1st Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, Prague, Czech Republic, pp. 13-20, 1957.
- [Eph] Y. Ephraim and N. Merhav, "Hidden Markov processes", *IEEE Transactions on Information Theory*, Vol. 48, No. 6, pp. 1518-1569, June 2000.
- [Tweem] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993.
- [SHME] S. Suvorova, S. D. Howard, W. Moran and R. J. Evans, "Waveform libraries for radar tracking applications: manoeuvring targets", *2005 Defence Applications of Signal Processing Workshop*, Midway UT, USA, March 2005.
- [WangMor] X. Wang and B. Moran, "Multitarget Tracking Using Virtual Measurement of Binary Sensor Networks", Submitted to *9th International Conference on Information Fusion*, Florence, Italy, July 2006.
- [WangMus] X. Wang and D. Mušicki, "Enhanced Tracking Performance with Signal Amplitude Information of Sensor Networks", Submitted to *9th International Conference on Information Fusion*, Florence, Italy, July 2006.
- [LaS] B. F. La Scala, "Scheduling Binary Proximity Sensor for Robust Target Tracking", Submitted to *9th International Conference on Information Fusion*, Florence, Italy, July 2006.
- [LaSMOREsav] B. F. La Scala, M. R. Morelande and C. O. Savage, "Robust Target Tracking with Unreliable Binary Proximity Sensors", *IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse France, June 2006.
- [MusEvans] D. Mušicki and R. J. Evans, "Measurement Gaussian Sum Mixture Target Tracking", Submitted to *9th International Conference on Information Fusion*, Florence, Italy, July 2006.
- [MusMore] D. Mušicki and M. R. Morelande, "Tracking a Large Number of Targets in Clutter with Particle Filter", Submitted to *14th European Signal Processing Conference*, Florence Italy, September 2006.
- [Ach] D. Achlioptas, "Database-friendly Random Projections", *20th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, Santa Barbara CA, USA, pp 274-281, 2001.

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

[JohnLin] W. B. Johnson and J. Lindenstrauss, "Extensions of Lipschitz Mappings into a Hilbert Space", *Conference in Modern Analysis and Probability*, New Haven CT, USA, pp 189-206, 1984.

[LaSMor] B. F. La Scala and B. Moran, "Measures of Effectiveness for Waveform Selection", *1st International Conference on Waveform Diversity and Design*, Edinburgh UK, November 2004.

[Reza] M. Rezaeian, "The entropy rate of the hidden Markov process", Submitted to *IEEE Transactions on Information Theory*, May 2004.

2. I. Acronyms

ADTS	Advanced Detection Technology Sensor
ASU	Arizona State University
ATA	Automatic Target Acquisition
AVU	Algorithms Verification Units
CADSP	Cooperative Analog Digital Signal Processor
CCDR	Classification Constrained Dimensionality Reduction
CRB	Cramér–Rao Bound
CROPS	Classification Reduction Optimal Policy Search
DARPA	Defense Advanced Research Projects Agency
DSA	Distinct Sensing Area
dwMDS	Distributed, weighted, multi-dimensional scaling
FPA	Focal Plane Array
FMAH	Fast Mathematical Algorithms and Hardware
GEM	Geometric Entropy Maps
Georgia Tech	Georgia Institute of Technology
GPS	Global Positioning System
IASG	Independently Activated Sensor Group
ISP	Integrated Sensing and Processing
IXO	Information Exploitation Office
kNN	k-Nearest Neighbor
LEAN	Laplacian Eigenmap Adaptive Neighbor
LIP	Linear Integer Programming
MC	Monte-Carlo
MTT	Multi-target tracking
NEST	Networked Embedded System Technology
NLIP	Nonlinear Integer Programming
NLOS	NetFires Non-Line of Sight
NUC	Non-Uniformity Compensation
ONR	Office of Naval Research
PAM	Precision Attack Munition
PDA	Probabilistic Data Association
PWF	Polarization Whitening Filter
PoP	Period of Performance
RIM	Radio Interferometric Measurements
RIPS	Radio Interferometric Positioning
RISCO	Raytheon International Support Company
RSS	Received Signal Strength

ISP Phase II (Contract N00014-04-C-0437)
Quarterly Progress Report (CDRL A001 No. 4)

TAA	Technical Assistance Agreement
TDOA	Time Difference of Arrival
TIM	Technical Interchange Meeting
UAV	Unmanned Aerial Vehicle
UCIR	Uncooled infrared imaging
UM	University of Michigan
UniMelb	Melbourne University
VM	Virtual Measurement
VU	Vanderbilt University